Joana Ferreira Guerreiro

Licenciada em Ciências e Engenharia do Ambiente

Assessment of NDVI, Land Surface Temperature and Precipitation anomalies for drought monitoring in Bayankhongor province, Mongolia.

Dissertação submetida para obter o Grau de Mestre em Engenharia do Ambiente, Perfil de Engenharia de Sistemas Ambientais

Orientadora: Prof. Doutora Maria Júlia Seixas, Professora Auxiliar, FCT/UNL

Co-orientador: Doutor Nuno Santos Grosso, DEIMOS Engenharia SA

Júri:

Presidente da mesa: Prof.ª Doutora Maria Teresa Calvão Rodrigues

Vogais: Doutor Nuno Miguel Matias Carvalhais
Prof.ª Doutora Maria Júlia Fonseca de Seixas

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Outubro 2015

Assessment of NDVI, Land Surface Temperature and Precipitation anomalies for drought monitoring in Bayankhongor province, Mongolia.

Acknowledgments

This thesis represents a milestone in six years of studying in FCT-UNL and I wouldn't been able to finish it without the guidance and support of several individuals.

First, I would like to thank my academic advisor, Professor Maria Júlia Seixas, for awakening my interest in remote sensing and being always available to answers my questions. Her mentoring, useful remarks and sympathy over the course of this experience helped me in all the time of writing this thesis.

I would also like to thank my co-advisor, Dr. Nuno Grosso, for the opportunity to work with a company outside the faculty which contributed to make my thesis experience richer and productive. All the encouragement through the learning process of this thesis was motivating and his patience and friendly personality through the hours devoted to finding and understanding errors in my python programs were greatly appreciated.

I acknowledge my gratitude to everyone in DEIMOS Engenharia, for receiving me with arms wide open which made my experience much more stimulating. A special thanks to Carla Patinha, for always being available to send me the necessary information with a smile on her face.

I take this opportunity to show my gratitude to my parents for their love and continuous support throughout my life. They always encouraged me and gave me the strength to reach my goals and chase my dreams. Thank you for your patience and comprehension during the long days in front of the computer when no one could speak or make any noise. You are my safe haven!

To my boyfriend, Fábio Santos, I wish to offer my deepest thanks. He stood by me cheering me up through the process, even when I retreated to long days in the company of my computer. You are my rock and your support and faith in me makes me a better person. You make my life fuller and I want to thank you for always being there when I need you the most.

I would like to thank my best friends, Rita and Katy, for giving a meaning to the expression "true friendship". Rita we've been through so much together in these 23 years. Despite going our separate ways we've always managed to be true to ourselves and stay close to one another. Thank you for always showing up in the most important days. Katy, you've become my partner in crime and my confidant. We are so different, yet so close and that's what makes our friendship so unique. Thank you for being there for me all these years smiling, studying, crying, laughing, talking, singing… being my friend.

Andreia and Afonso, I want to thank you for your friendship and support. My life has become happier and funnier since that first math class. Sandra and Carol, thank you for being the proof that high school friends can stay in our lives and support each other through the years.

Finally, I would like to thank my family and all the friends who supported me and made the writing of my thesis a more enjoyable and less stressful process. Thank you all!

Abstract

During the last decade Mongolia's region was characterized by a rapid increase of both severity and frequency of drought events, leading to pasture reduction. Drought monitoring and assessment plays an important role in the region's early warning systems as a way to mitigate the negative impacts in social, economic and environmental sectors. Nowadays it is possible to access information related to the hydrologic cycle through remote sensing, which provides a continuous monitoring of variables over very large areas where the weather stations are sparse.

The present thesis aimed to explore the possibility of using NDVI as a potential drought indicator by studying anomaly patterns and correlations with other two climate variables, LST and precipitation. The study covered the growing season (March to September) of a fifteen year period, between 2000 and 2014, for Bayankhongor province in southwest Mongolia. The datasets used were MODIS NDVI, LST and TRMM Precipitation, which processing and analysis was supported by QGIS software and Python programming language. Monthly anomaly correlations between NDVI-LST and NDVI-Precipitation were generated as well as temporal correlations for the growing season for known drought years (2001, 2002 and 2009).

The results show that the three variables follow a seasonal pattern expected for a northern hemisphere region, with occurrence of the rainy season in the summer months. The values of both NDVI and precipitation are remarkably low while LST values are high, which is explained by the region's climate and ecosystems. The NDVI average, generally, reached higher values with high precipitation values and low LST values. The year of 2001 was the driest year of the time-series, while 2003 was the wet year with healthier vegetation.

Monthly correlations registered weak results with low significance, with exception of NDVI-LST and NDVI-Precipitation correlations for June, July and August of 2002. The temporal correlations for the growing season also revealed weak results. The overall relationship between the variables anomalies showed weak correlation results with low significance, which suggests that an accurate answer for predicting drought using the relation between NDVI, LST and Precipitation cannot be given. Additional research should take place in order to achieve more conclusive results. However the NDVI anomaly images show that NDVI is a suitable drought index for Bayankhongor province.

Keywords: Anomaly, Correlation, Drought Indicator, LST, NDVI, Precipitation, Remote Sensing

Resumo

Na última década, a região da Mongólia foi caracterizada por um rápido aumento da frequência e severidade de eventos de seca, levando à redução de pastagens. A avaliação e monitorização da seca tem um papel importante nos sistemas de alerta da região como uma forma de mitigar os impactes negativos nos sectores social, económico e ambiental. Indicadores e índices de seca são utilizados para caracterizar a severidade, dimensão espacial e duração das secas. Atualmente é possível aceder a informação relacionada com o ciclo hidrológico através da deteção remota, o que permite uma monitorização contínua de variáveis ao longo de áreas maiores onde estações meteorológicas não tem uma expressão tão significativa.

A presente tese teve como objetivo, explorar a possibilidade de utilizar o NDVI como potencial indicador de seca, através do estudo de padrões de anomalias e correlações com outras duas variáveis climáticas (LST e precipitação). O estudo teve como dimensão temporal o período de crescimento vegetativo (Março a Setembro) de um intervalo de tempo de quinze anos (2000-2014) para a província de Bayankhongor no sudueste da Mongólia. Os *data sets* utilizados foram NDVI e LST do MODIS e a Precipitação da TRMM, cujo processamento e análise foi suportado pelo software QGIS e a linguagem de programação Python. Foram calculadas correlações mensais de anomalias entre NDVI-LST e NDVI-precipitação, assim como correlações temporais para o período de crescimento vegetativo para três anos de seca conhecidos (2001,2002 e 2009).

Os resultados mostraram que as três variáveis seguem um padrão sazonal expectável para uma região do hemisfério norte, cujo período de chuvas ocorre nos meses de Verão. Os valores de NDVI e precipitação são notavelmente baixos enquanto os valores de LST são elevados, o que é explicado pelo clima e ecossistemas da região. O ano de 2001 foi o mais seco da série temporal e 2003 foi o ano com maior crescimento vegetativo. No geral, a média de NDVI atingiu valores mais elevados quando os valores de precipitação eram também mais elevador e os valores de LST mais baixos, o que evidencia uma possível correlação positiva com a precipitação e uma correlação negativa com a LST.

As correlações mensais registaram resultados fracos e com pouca significância, à exceção das correlações entre NDVI-LST e NDVI-precipitação para Junho, Julho e Agosto de 2002. As correlações temporais para o período de crescimento vegetativo e para a combinação de Junho, Julho e Agosto, também revelaram resultados fracos. No geral, as relações entre as anomalias das três variáveis em estudo apresentaram baixas correlações com resultados pouco significativos, o que sugere que a previsão de secas baseada na relação entre NDVI, LST e a Precipitação ainda está por determinar. Estudos adicionais deverão ser levados a cabo de forma a obter resultados mais fiáveis e conclusivos. Contudo, as imagens obtidas para as anomalias de NDVI sugerem que este é um bom indicador de seca para a província de Bayankhongor.

Palavras-chave: Anomalia, Correlação, Deteção Remota, Indicador de Seca, LST, NDVI, Precipitação

Table of Contents

List of Figures

List of Tables

Acronyms

AVHRR - Advanced Very High Resolution Radiometer

EO – Earth Observation

EPSG - European Petroleum Survey Group

ET – Evapotranspiration

GEO – Geostationary

GIS – Geographic Information System

LEO - Low Earth Orbit

LST – Land Surface Temperature

MODIS - Moderate Resolution Imaging Spectroradiometer

NDVI – Normalized Difference Vegetation Index

QC - Quality Control

SDS – Scientific Data Sets

TRMM - Tropical Rainfall Measuring Mission

UTM - Universal Transverse Mercator

VCI – Vegetation Condition Index

WGS - World Geodesic System

1. Introduction

Extreme climatic events have been increasing in frequency and severity worldwide over the past century (John, et al., 2013). Drought is possibly the most complex and harmful natural hazard since it shows a highly variability over time and space and usually extends to large geographic areas (Senay, et al., 2015). Over the past five decades the global land area affected by drought has increased, particularly in central and northern Eurasia (Zhang, et al., 2012).

This kind of extreme event carries a handful of impacts that span various sectors of the society causing dire environmental, social and economic consequences. Drought is the number one natural hazard when speaking about the number of people affected (Mishra & Singh, 2010). Therefore, assessment and monitoring of droughts is crucial to early warning systems and management in order to mitigate drought's serious effects.

Efforts have been made to develop and implement various quantitative measures of drought's extent and severity. However, it is difficult to predict and monitor using traditional approaches, especially over large areas. The use of satellite remote sensing in drought monitoring has gained more attention since it can be used to measure meteorological or biophysical characteristics of Earth's surfaces (Rhee, Im, & Carbone, 2010). Technological advances in this field have enabled continuous data measurements over a range of spatial and temporal scales which can help generate long term information on drought events (Senay, et al., 2015).

Remote sensing-based vegetation indices have been widely used for drought monitoring and tracking. Among the vegetation indices, the Normalized Difference Vegetation Index (NDVI) is the most common. The vegetation condition index (VCI) also measures the time of drought's onset and its intensity, duration and impact on vegetation. However, it is best used during growing season when vegetation is most photosynthetically active (Mishra & Singh, 2010). Also, the combination of vegetation indices with drought indices and data such as land surface temperature and rainfall can be a very powerful instrument as it provides useful and more detailed information on drought monitoring ((Karnieli, et al., 2010), (Nichol & Abbas, 2015)).

Mongolia, already characterized by severe weather conditions due to its inland location, is one of the regions dealing with extreme climatic events caused by climate change impacts. These events include droughts, *dzuds* (harsh winters), dust storms and desertification. During the last decade the severity and frequency of drought during the growing season has increased as well as extreme winters, leading to pasture reduction in Mongolia which is a key economic drive of the country (Nandintsetseg & Shinoda, 2012). The most recent extreme events in Mongolia were the combined summer drought–*dzud* events of 2000–2002 and the drought-*dzud* event of *2009*- 2010 (John, et al., 2013). Accordingly, an early warning system and a better understanding of summer drought events is a main concern for this country in order to mitigate drought's impacts on the economy and human subsistence.

1.1. Motivation, Scope and Objectives

The motivation for this dissertation arose from the opportunity of working in satellite remote sensing with DEIMOS Engenharia, who is developing a project for the Asian Development Bank. This project, “Climate-Resilient Rural Livelihoods in Mongolia”, aims to complement “government efforts to develop a sustainable, climate proof livestock sector, the main economic activity in Mongolia, able to overcome the productivity and in-come loss problems, related to over-grazing and climate change, registered in recent years”. The project delivers two main services: (1) Land use/land cover mapping and Digital Elevation Model; (2) Drought monitoring, focused to enhance the current National Remote Sensing Centre drought monitoring system which is based on Moderate Resolution Imaging Spectroradiometer (MODIS) data. The drought monitoring service is based on NDVI data applied to Bayankhongor province in Mongolia (DEIMOS Engenharia, 2015).

The second service being rendered by DEIMOS was the main drive for defining the purpose and scope of the thesis. Thus, the aim of the study is to explore the possibility of using NDVI anomalies as a drought index, based on its relation with land surface temperature (LST) and precipitation data. The study site refers to Bayankhongor province, in southwest Mongolia, and is based in the growing season (March to September) of a fifteen year time series (2000-2014), with a focus on known drought years (2001, 2002 and 2009). The analysis was supported by QGIS software and Python programming language. Several objectives were defined to lead to the main goal of the study:

i) Calculation of NDVI, LST and Precipitation monthly anomalies for the time period 2000-2014;
ii) Assessment of correlation between NDVI, LST and Precipitation anomalies for the drought years 2001, 2002 and 2009:
    - Calculation of monthly correlations;
    - Calculation of growing season temporal correlations.
iii) Exploratory analysis and interpretation of the correlation patterns between NDVI and LST and Precipitation, by using land use and land cover maps.

1.2. Thesis outline

This document consists in five chapters, including the Introduction (this chapter), with specific focus and objectives:

Chapter 2 - Literature Review: It describes and analyses previous research related to the topic as well as defines important concepts. This chapter covers three main areas related to the problem, (1) drought concepts and drought monitoring challenges, (2) using NDVI as a drought monitor index, and (3) additional datasets used as complement to the NDVI.

Chapter 3 - Materials and Methods: This chapter describes the data being used and outlines the procedures used to conduct the study. It includes four main components, (1) description of Bayankhongor province characteristics, (2) description of the three datasets used for the study (NDVI, LST and Precipitation), (3) data pre-processing methods as a first approach to the datasets, and (4) data processing which describes the methodology used to generate results, specifically anomalies calculations, monthly anomaly correlations between datasets and final time series spatial correlation.

Chapter 4 - Results and Discussion: This chapter reports and addresses the results generated from the previous chapter and discusses the significant findings and its meaning for the study.

Chapter 5 - Conclusions: Identifies the critical conclusions gathered from the results and findings discussion and relates them with the research objectives defined in the Introduction. It includes a *Limitations* section, where are presented limitations encountered during the research process, and a *Further Investigation* section, where it's mentioned future research and follow-up studies related to the present study.

2. Literature Review

2.1. Drought concepts

Drought is a complex natural hazard that carries negative impacts for people and for the environment. It is typically characterized by type, frequency, duration, magnitude, severity and geographic extent. However, determining a universal definition is extremely difficult due to drought's temporal and spatial variation. A general definition of drought defines it as "*an extended period – a season, a year, or several years – of deficient rainfall relative to the statistical multi-year average for a region*" (Graham, 2000)

There are four main types of droughts defined in the literature (Senay, et al., 2015):

I. **Meteorological drought,** which is defined by a lower precipitation than the long-term average precipitation, for a prolonged period of time (i.e. it is based on the degree of dryness and the duration of that dry period);
II. **Agricultural drought,** happens when the available water for plants and crops falls below the required limit not meeting the water needs of that specific crops and, thus, limiting vegetation's growth; it usually happens after a meteorological drought and before a hydrologic drought;
III. **Hydrologic drought,** is associated with the lack of availability of surface and subsurface water supplies (stream flow, soil moisture and groundwater);
IV. **Socio-economic drought**, measured by social and economic indicators, is based on the impact of meteorological, agricultural and hydrological drought on supply and demand of economic goods; it happens when a climate related deficit in water supplies results in demand exceeding the supply.

The most common types mentioned in remote sensing studies are meteorological drought and agricultural drought. For the purpose of this thesis the definition of drought used is agricultural drought.

The reasons behind a drought are intricate because they depend, not only, on the atmosphere but also on the hydrologic processes that deliver moisture to the atmosphere. When dry hydrologic conditions are established, the reduced moisture in the soil's upper layers translates into a decrease in the evapotranspiration rates which leads to a reduction of the atmosphere's relative humidity (Mishra & Singh, 2010). The less humidity there is the less probable it becomes to occur precipitation which aggravates the lack of moisture in the soil, affecting plant growth over time (Ji & Peters, 2003). Thus, all types of droughts can be associated with precipitation deficit and each element of the hydrologic cycle have a different response to drought events (AghaKouchak, et al., 2015).

2.2. Drought in the Mongolia's region

Mongolia is located in Central Asia with a total area of 1 564 116 km^2 (Erch Partners, 2014), bounded on the north by Russian Federation and on the east, south and west by People's Republic of China. The country's topography consists mainly of steppe[1] with mountain ranges in the north and west areas. Most of the territory is considered arid, semi-arid, moderate arid and moisture deficient regions. The Gobi desert occupies 41,3% of Mongolian territory which makes drought and desertification an important issue to discuss (UNEP, 2002). Mongolia landlocked location combined with sparsely population and harsh climate results in a quite vulnerable country to changing weather conditions (Erch Partners, 2014).

Having livestock breeding has a key traditional economic sector, Mongolia's population and its pastoral activities are susceptible to the recurring drought events from the last decades. The pasture quality and resources become unpredictable and the weather limitations for agricultural production in the steppe area are a threat for the sustenance of most population. In addition there is the danger of drought aggravating extreme winter conditions which can lead to losses of livestock (Sternberg, Thomas, & Middleton, 2011).

Since 1940, Mongolia's annual average temperature has suffered a 1,9°C rise while the annual precipitation decreased until the mid-1980s. Precipitation has been increasing ever since, except in the Gobi desert area. (Batima, Natsagdorj, Gombluudev, & Erdenetseteg, 2005). The rise of temperature alongside with precipitation fluctuations, as a result of global climate trends, have created a situation where the weather inconsistency and extreme events work together to aggravate drought occurrence in the region (Sternberg, Thomas, & Middleton, 2011).

During the 2000s growing season droughts have increased in both frequency and severity across Mongolia's region. A study from Nandintsetseg and Shinoda (2012) have shown that consecutive droughts during the growing season in Mongolia contributed to a reduction of pasture production, especially during June-August of 2000-2002 and 2007. During the last decade the most extreme events on vegetation were the summer drought–*dzud* events of 2000–2002 and the drought-*dzud* event of *2009*- 2010 (John, et al., 2013). The study carried from John et al (2013) identified 2000-2001, 2005 and 2009 as particularly dry years in the desert biome, which revealed to be more vulnerable to drought than the grassland biome.

The rapid increase of frequency and severity of droughts in this region have been alarming during the last decade (John, et al., 2013). The assessment of this extreme event and its monitoring should become a priority in order to study efficient warning systems that can help mitigate drought's negative impacts in Mongolia's social, economic and environmental sectors.

[1] Semi-arid region consisting of a dry, grassy plain that occurs in temperate climates. The Eurasian steppe, extending from Hungary to China, is the largest temperate grassland in the world.

1.1. Drought Monitoring through Earth Observation data

The detection and monitoring of drought presents a challenge since these events are difficult to quantify and define in time and space. First, droughts develop slowly and their effects can increase gradually and often accumulate and linger over a considerable period of time. This makes it difficult to determine the onset and end of a drought. Second, the concept of drought doesn't have a precise and universal definition which leads to confusion. Third, drought's impacts are non-structural and usually spread over large geographical areas. Fourth, contrary to other natural hazard droughts can be directly triggered by anthropogenic activities such as over farming, deforestation, excessive irrigation and erosion (Mishra & Singh, 2010).

Despite the challenges, efforts have been made to study and gain insight on this phenomenon through the establishment of drought indicators and indices and analysis of earth observation data. Drought indicators and indices are different concepts broadly used to characterize drought's severity, spatial extend and duration. While an *indicator* consists of parameters (precipitation, temperature, streamflow, groundwater levels, reservoir levels, soil moisture levels, snow pack and drought indices), an *indice* often refers to a combination of indicators which outcome is a computed numerical value of a drought's severity or magnitude (Wardlow, Anderson, & Verdin, 2012).

As a result of technological advances, it is now possible to obtain an increasing number of variables related to the hydrologic cycle (Wardlow, Anderson, & Verdin, 2012). These include: precipitation, vegetation condition, soil moisture, groundwater and evapotranspiration (ET). Both precipitation and vegetation condition can be directly estimated from remotely sensed data while the other parameters have to be modelled (Senay, et al., 2015). Each variable is converted into a drought indicator through the calculation of an anomaly's extent, having a long term series as a baseline (AghaKouchak, et al., 2015). They can then be used to assess drought's severity, showing the incredible potential of remote sensing contribution to drought monitoring (Wardlow, Anderson, & Verdin, 2012).

There are two types of remote sensing satellites broadly used in drought monitoring and impact assessment: (1) Geostationary (GEO) satellites, that by being synchronized with Earth's rotation provide a constant view of the same surface area which is very useful for weather monitoring; (2) Low Earth Orbit (LEO) satellite, that have a Sun-synchronous orbit[2] allowing more than one image per day and a comparison of images without extreme changes in shadows and lighting (Riebeek, 2009). While GEO satellites carry multispectral radiometers that collect data in the visible and infrared portion of the electromagnetic spectrum, LEO satellites carry a variety of sensors, such as multispectral and hyperspectral sensors, laser altimeters, microwave sensors, among others (AghaKouchak, et al., 2015).

[2] Crosses over the equator at approximately the same local time, on the ground, each day (and night) keeping the sunlight's angle as consistent as possible (Riebeek, 2009) .

The earnest use of satellite remote sensing data regarding Earth's weather and climate started in 1960 with Television and Infrared Observation Satellite (TIROS-1) mission, which success led to an increase of this sort of missions. Figure 2.1 shows a rough list of current and future satellite sensors and missions that are relevant to drought monitoring. Some of the most relevant being the Global Precipitation Mission (GPM), Geostationary Operational Environmental Satellites R series (GOES-R), GRACE Follow-On, SMAP, and SWOT missions (AghaKouchak, et al., 2015).

Both precipitation and soil moisture are essential components of the water cycle and have been used for drought monitoring and prediction. A variety of techniques have been developed for routine retrieval of rainfall using satellite data, which can then be used to calculate indices such as the Standardizer Precipitation Index (SPI) to measure meteorological drought. Some of the available precipitation data sets include the Climate Predicting Centre (CPC) Morphing Technique (CMORPH), Tropical Rainfall Measuring Mission (TRMM), Multi-satellite Precipitation Analysis (TMPA), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) and the Global Precipitation Climatology Project (GPCP). The Climate Change Initiative (CCI) for Soil Moisture provides soil moisture data derived from multiple satellite-based sensors (AghaKouchak, et al., 2015).

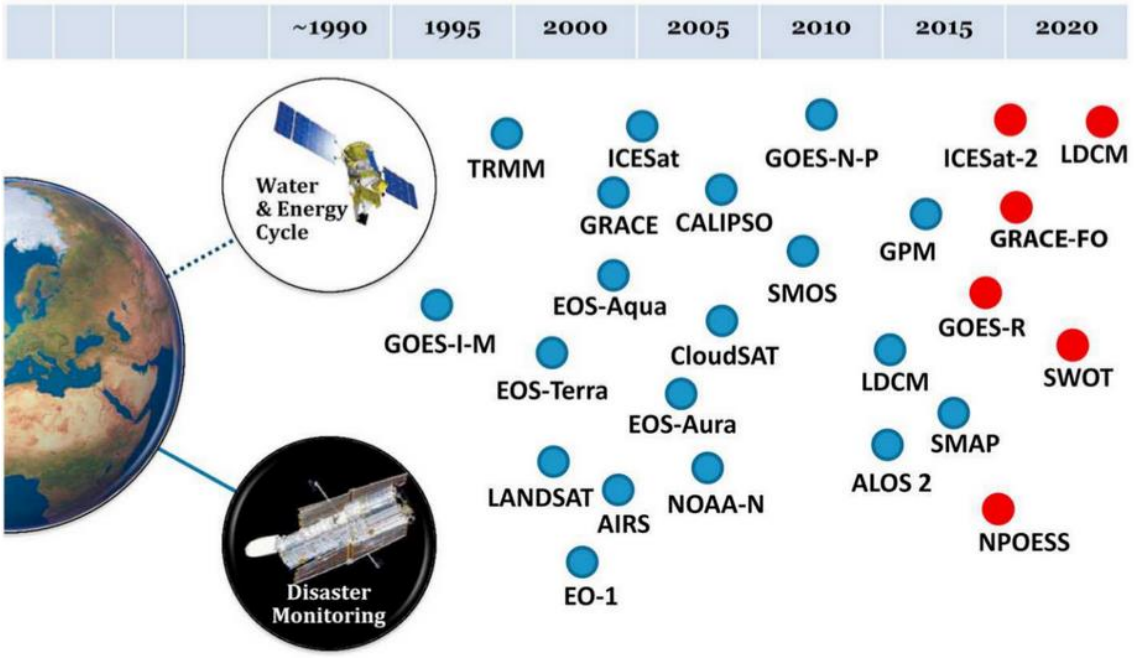


Figure 2.1 | Current and future satellite missions relevant to drought monitoring and assessment

(TRMM: Tropical Rainfall Measuring Mission; GRACE: Gravity Recovery and Climate Experiment; FO: Follow-On; ICESat: Ice, Clouds, and Land Elevation Satellite; CALIPSO: Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations; EOS-Aqua: Earth Observing System Aqua; EOS-Terra: Earth Observing System Terra; EOS-Aura: Earth Observing System Aura; AIRS: Atmospheric Infrared Sounder; EO-1: Earth Observing-1; GOES: Geostationary Operational Environmental Satellite; NOAA-N: NOAA Polar Operational Environmental Satellites N Series; SMOS: Soil Moisture and Ocean Salinity satellite;ICESat-2: Ice, Clouds, and Land Elevation Satellite; GPM: Global Precipitation Measurement; LDCM: Landsat Data Continuity Mission; SWOT: Surface Water and Ocean Topography; SMAP: Soil Moisture Active Passive; ALOS 2: Advanced Land Observing Satellite; and NPOESS: National Polar-orbiting Operational Environmental Satellite System) (AghaKouchak, et al., 2015)**.**

Evapotranspiration (ET) is also an important component of the water and energy cycle and plays an important role in drought monitoring. Evapotranspiration describes water/moisture availability as well as the rate at which it is consumed in ecosystems. Some of the drought indicators being used that integrate ET data are: WaterStress Index (CWSI), Water Deficit Index (WDI), Evaporative Stress Index (ESI), Evaporative Drought Index (EDI), Drought Severity Index (DSI) and Reconnaissance Drought Index (RDI) (AghaKouchak, et al., 2015).

Droughts are naturally associated with vegetation conditions and cover, which enabled the application and study of vegetation indices (VI) for drought assessment (Karnieli, et al., 2010). Satellite-based remote sensing suffered a dramatic transformation with the launch of National Oceanic and Atmospheric Administration Advanced Very High Resolution Radiometer (NOAA AVHRR) instrument in 1979 (Wardlow, Anderson, & Verdin, 2012). This instrument provided systematic monitoring of vegetation patterns and conditions, allowing the application of NDVI data to drought monitoring (AghaKouchak, et al., 2015). The use of AVHRR vegetation derived data has several advantages over meteorological drought indices. AVHRR data has a higher spatial density, 1km pixel, of data collection when compared to weather stations and its sensor covers very large areas. Regions that present a low density of weather stations can still have available data from AVHRR (Ji & Peters, 2003).

During the last decade, studies have shown that the integration of multiple indices and data sets improves drought assessment. Thus, drought monitoring should be based on multiple variables in order to provide a more robust and cohesive measure of drought that captures the diverse range of vegetation response to drought across different ecosystems (AghaKouchak, et al., 2015). There are still major challenges including data continuity, unquantified uncertainty, sensor changes, global community acceptability and data maintenance (AghaKouchak, et al., 2015). However, considering the limitations associated with ground-based methods, remote sensing will continue to complement the traditional methods while evolving through technologic advances and gaining worldwide acceptance. As this study area progresses, drought assessment will move forward resulting in better monitoring at multiple spatial scales (Wardlow, Anderson, & Verdin, 2012).

2.3. NDVI as a drought indicator

In the late seventies, the relationship between photosynthesis and the amount of photosynthetic active radiation absorbed by plants was discovered. The more solar radiation a plant absorbs the more productive that plant is going to be due to the high rate of photosynthesis. This happens to a certain extent, when other limiting factor take place and have an impact in the plant's productivity. Since then, the observation of this relationship has allowed to create normal patterns of plants growing conditions for a certain region for a given time of the year (Weier & Herring, 2000).

The Normalized Difference Vegetation Index (NDVI) (Tucker, 1979), computed from satellite AHVRR radiance data, is the most commonly used vegetation index and is mathematically represented by the following formula:

$$NDVI = \frac{NIR-VIS}{NIR+VIS} \qquad (1)$$

Where NIR refers to near-infrared and VIS to visible light. The outcome of this expression ranges from -1 to +1. A value of zero translates into no vegetation while a value close to +1 points to the highest possible density of green leaves. As a result, a given site's absorption and reflection of photosynthetically active radiation over a determined period of time can be used to describe vegetation's health conditions in that site, relative to the norm (Weier & Herring, 2000).

This index has drawn attention for its use in drought monitoring worldwide considering that vegetation's density is closely related to land surface moisture conditions. Nevertheless it is important to mention that it's most effective when takes in account the seasonality and it's used during growing season when vegetation shows a higher primary production. Whereas NDVI shows promising results during growing season, its utility is limited during cold season due to the vegetation's dormancy (Ji & Peters, 2003). An early study from 1987, concluded that NDVI data could be used to identify and quantify droughts in semiarid and arid regions (Karnieli, et al., 2010).

The most common and simplest methods regarding NDVI data for drought assessment use NDVI's anomalies. An anomaly in NDVI data is detected by calculating the difference between the NDVI composite of a certain time period and the long term mean NDVI for the same period using several years of data record. While a positive anomaly indicates a vegetation growth above the normal vegetation condition, a negative anomaly can indicate a drought situation (Weier & Herring, 2000). This way it is possible to isolate the vegetation signal variability and establish a historical context for the current NDVI, which is a more accurate description over larger areas and a more intuitive analysis.

In the Sahel region (Africa), a study showed that negative NDVI anomalies could identify the spatial extent of drought response in vegetation (Wardlow, Anderson, & Verdin, 2012). However, NDVI anomalies can be caused by a variety of events non-related to drought. Fire, land cover change, plant disease, pest infestation, biomass harvesting and flooding can generate anomalies similar to those caused by drought. A more consistent analysis and reliable results can be achieved by using additional indicators to complement satellite-based NDVI (AghaKouchak, et al., 2015).

A study undertaken by John *et al*. (2013) in the Mongolian plateau during the period from 2000 to 2010 mapped vegetation indices and land surface temperature anomalies to assess vegetation response to extreme climate events. The authors concluded that drought events substantially reduced vegetation activity in that region, which is an indicator that vegetation indices can be a useful to assess drought events. In recent years, numerous ways of combining both NDVI and

LST data have been explored for drought monitoring. Several studies ((Son, Chen, Chen, Chang, & Minh, 2012), (Sruthi & Mohammed Aslam, 2015), (Karnieli, et al., 2010)) have shown that the assessment of both NDVI and LST data can provide information on vegetation and moisture conditions which leads very useful insight regarding agricultural drought monitoring. Sruthi and Aslam (2015) studied the comparison between NDVI and LST in an Indian region prone to drought. Their findings revealed a high negative correlation between the two datasets and concluded that the studied combination can be used to detect agricultural drought.

The combination of NDVI and LST data for drought assessment depends on seasonality and time of the day. In regions where the water is the limiting factor for vegetation growth the correlation between NDVI and LST is negative, while in regions where the solar radiation is the limiting factor for vegetation growth a positive correlation exists between the two variables (Karnieli, et al., 2010). According to Karnieli (2010), it is recommended to assess the relationships between NDVI and LST for drought monitoring in regions where water is the primary limiting factor.

A significant relationship has been described between NDVI and precipitation and soil moisture. As a result, NDVI has been widely used to address drought monitoring (AghaKouchak, et al., 2015). A study conducted by Ji and Peters (2003) in the U.S. northern Great Plains, studied the characteristics of the relationships between NDVI and the Standardized Precipitation Index (SPI), a meteorologically based drought index. The findings of the study suggested the effectiveness of NDVI, based on high correlations during the growing season.

3. Materials and Methods

3.1. Study Area

The study area was defined under DEIMOS's project and refers to the Bayankhongor aimag (province), which lies in southwest Mongolia, region of North East Asia (Figure 3.1 | . The province capital is also named Bayankhongor and has the highest population density of the province. Bayankhongor has a territory of 115 977.80 km^2 and a total population of 82 884 (ХЭЛТЭС, БАЯНХОНГОР АЙМАГ СТАТИСТИКИЙН, 2014).

Mongolian climate comprehends four seasons characterized by high fluctuations in temperature and low annual precipitation, as well as an average of 260 annual sunny days and cold winters. The severe climate is due to the great distance from the oceans, high elevation and the high mountain ranges surrounding the country (The World Bank Group, 2015). The growing season[3] ranges from March to September and 85% of total annual precipitation takes place during the summer months - June, July and August. The average temperature in Bayankhongor ranges from 0 to 7°C at the northern area and reaches an average of 8°C at the south and low regions. It can reach maximum temperatures from 28°C to 49° C and minimum temperatures at the Khangai Mountains can drop to -30°C (Erch Partners, 2014).

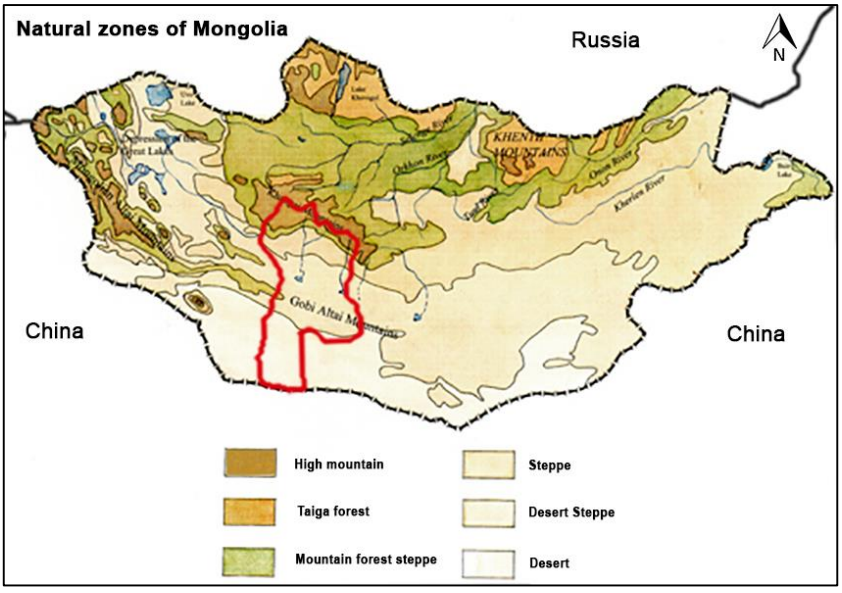


Figure 3.1 | Bayankhongor province location (red outline) and Mongolia's natural zones. (Adapted from Finch, C.(ed). Mongolia's wild heritage: Biological Diversity, Protected Areas, and Conservation in the Land of Chingis Khan. Boulder, CO: Avery Press, 1999.)

Bayankhongor province includes a variety of geographic areas that can be summed up into five regions: Khangai Mountain Ranges covered with taiga forest, mountain forest steppe, steppe, desert steppe and Gobi desert. The majority of the territory is covered by desert and desert steppe

[3] Period of the year when crops and other plants grow successfully.

area (Figure 3.1) with sparsely vegetation, mainly shrubs and low grass, and a hard mix of sand, clay and breakstone. The precipitation in this area varies throughout the province, decreasing from north to south. The forest steppe zone has an annual precipitation over 250 mm, the mountain forest steppe and steppe reach precipitation values between 130 and 180 mm, the desert steppe ranges from 50 to 100 mm and the desert areas have values between 0 and 50 mm (Erch Partners, 2014).

3.2. Datasets Description

3.2.1. NDVI

The NDVI dataset was acquired from MOD13Q1 product, Vegetation Indices 16-Day L3 Global 250m. The primary goal of this product is to provide comparisons of vegetation conditions every sixteen days with a spatial resolution of 250 m. It has a temporal coverage from February 2000 to the present time and presents a Sinusoidal projection. MOD13Q1's NDVI is computed from bi-directional surface reflectances atmospherically corrected and masked for water, clouds, heavy aerosols, and cloud shadows (LP DAAC, 2014).

1.1.1. Land Surface Temperature

For the Land Surface Temperature (LST) data it was used the MODIS product MOD11A2, level-3 MODIS global Land Surface Temperature (LST) and Emissivity 8-day. This dataset is composed from the daily 1-km LST product (MOD11A1) and represents average values of clear sky LSTs with a temporal resolution of eight days and a temporal coverage from March 2000 to the present time. The LST data set has a spatial resolution of 0,01°(1 km) and is archived in Hierarchical Data Format - Earth Observing System (HDF-EOS) format files. Each HDF LST file has multiple sub datasets (SDS) within. For MOD11A2 the relevant layers for the study are daytime LST and quality control (QC) (Land Processes Distributed Active Archive Center (LP DAAC), 2014).

3.2.2. Precipitation

Precipitation data was acquired from the Tropical Rainfall Measuring Mission Project product 3B43, which algorithm merges satellite and *in situ* data. It uses multiple independent precipitation estimates from satellites and monthly accumulated rain gauges *in situ* compiled by the Global Precipitation Climatology Centre (GPCC). The final product sums 3-hourly multi-satellite fields for each month and combines them with the monthly gauge analysis, representing the best estimate precipitation rate (mm/hr).

The temporal coverage of the data set is from January 1998 to April 2015 while the spatial coverage ranges from 50° South to 50°North latitude. It has a monthly temporal resolution and a spatial resolution of 0,25° (pixels of approximately 27km by 27 km) (Tropical Rainfall Measuring Mission Project (TRMM), 2011). The Table 3-1 shows a summary of each dataset essential information.

Table 3-1 | Summary of the data sets.

Variable	Dataset	Temporal Coverage	Temporal Resolution	Spatial Resolution	Format
NDVI	MOD13Q1	February 2000 – present	Sixteen days	250 m	HDF - EOS
Land Surface Temperature	MOD11A2.005	March 2000 - present	Eight days	0,01º (1 km)	HDF - EOS
Precipitation	TRMM_3B43	January 1998 - April 2015	Monthly	0,25° (27 km)	netCDF

3.3. Methodology

The Figure 3.3 represents the workflow of the methodology used in the present study. In the following sections each step of the methodology will be discussed.

3.3.1. Data pre-processing

The pre-processing of the data sets was an essential part of the study since the initial data sets were in different formats and projections and needed to be prepared for integration and analysis. It is important to mention that due to the large amount of initial files (approximately 2000), QGIS software[4] wasn't used to process each data set. Instead, the files were batch processed using Python[5] programs that were created for that purpose.

The study covers a fifteen year period from 2000 to 2014, focusing in the growing season months (March to September). The relation in time between the three datasets throughout the growing season can be seen in Figure 3.2. Both NDVI and LST datasets have a temporal resolution based on the Julian calendar, while Precipitation's follows the Gregorian calendar.

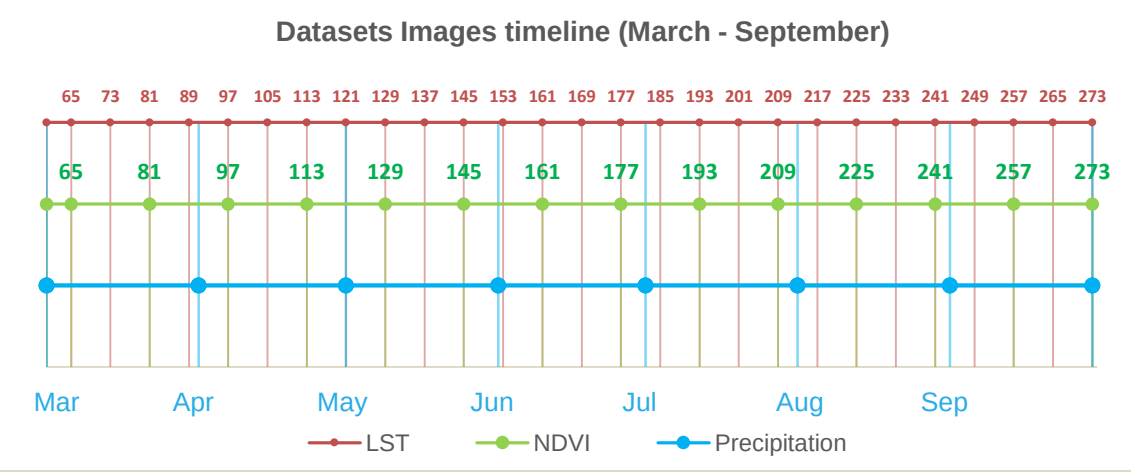


Figure 3.2 | Growing season datasets timeline with LST and NDVI following the Julian calendar.

[4] Open Source Geographic Information System software.
[5] Programming language available under an open source license.

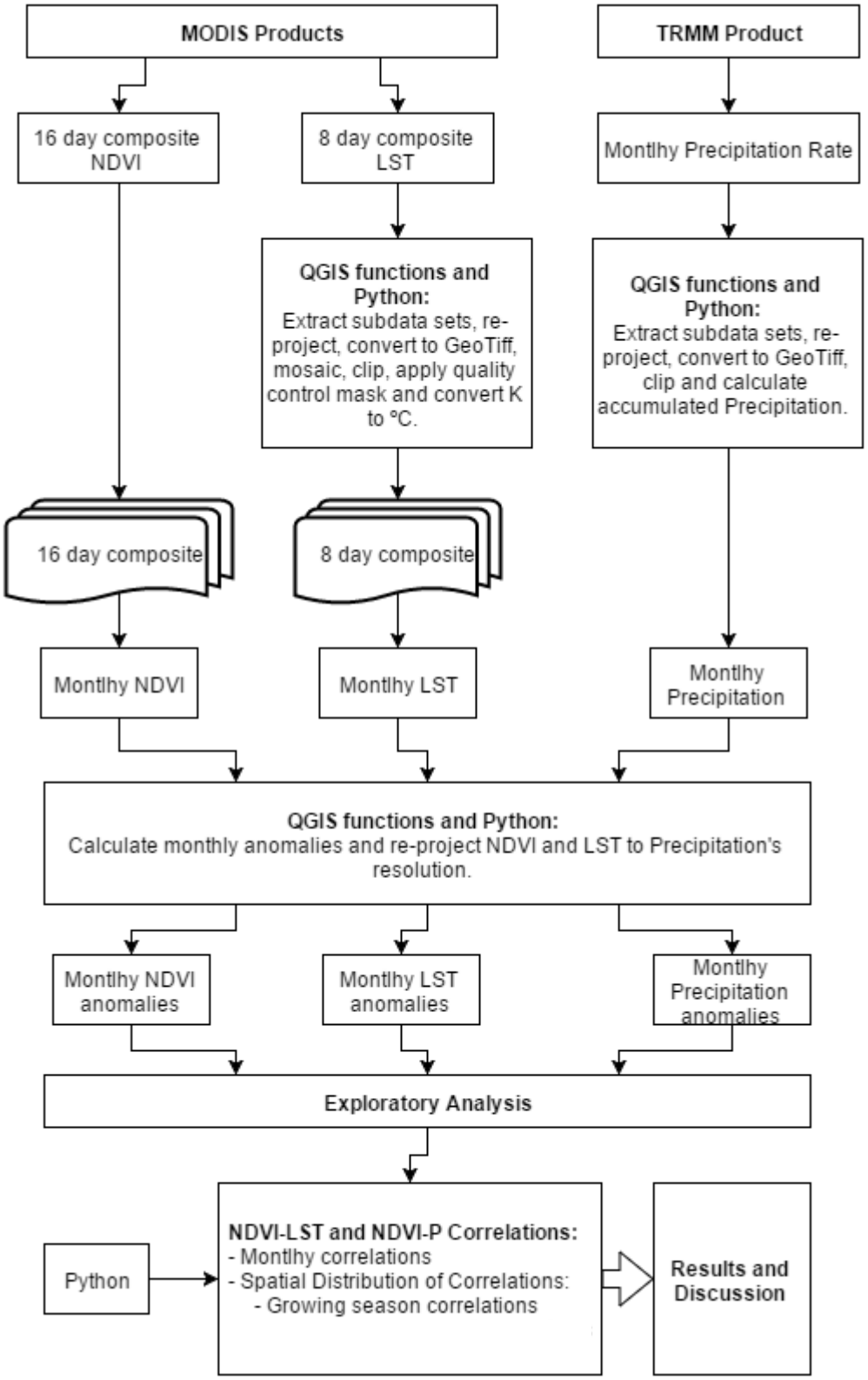


Figure 3.3 | Schematic representation of the methodology used

3.3.1.1. NDVI

The NDVI dataset was initially prepared by DEIMOS Engenharia's project team. All the files received were already in GeoTIFF format and in the EPSG 32647[6] projected coordinate systems. The 16-day images were filtered with a QC mask and clipped around Bayankhongor province.

Monthly NDVI composites were generated by averaging 16-day images corresponding to each month of the growing season during fifteen years. First, all the 16-day images were sorted into monthly folders, each folder containing the images from 2000 to 2014 for that specific month. Second, the average composites were calculated based on the pixel by pixel formula:

$$\boldsymbol{NDVI\,my = mean\,(NDVIy1, NDVIy2, \ldots NDVIyn)} \qquad \textbf{Equation 3-1}$$

Where:

- m, month
- y, year
- 1,2,… n, 16-day image (period).

The formula was applied by means of the Python program in Appendix A.5 to each monthly folder, year by year (Figure 3.4). In order to treat the information in Excel software and plot graphics, the program in Appendix A.10 was used. This program turns each image into an array of values and exports it to a csv file in a column format. Lastly, the NDVI monthly average for the fifteen year period was calculated by using again the program in Appendix A.5 with different file name patterns.

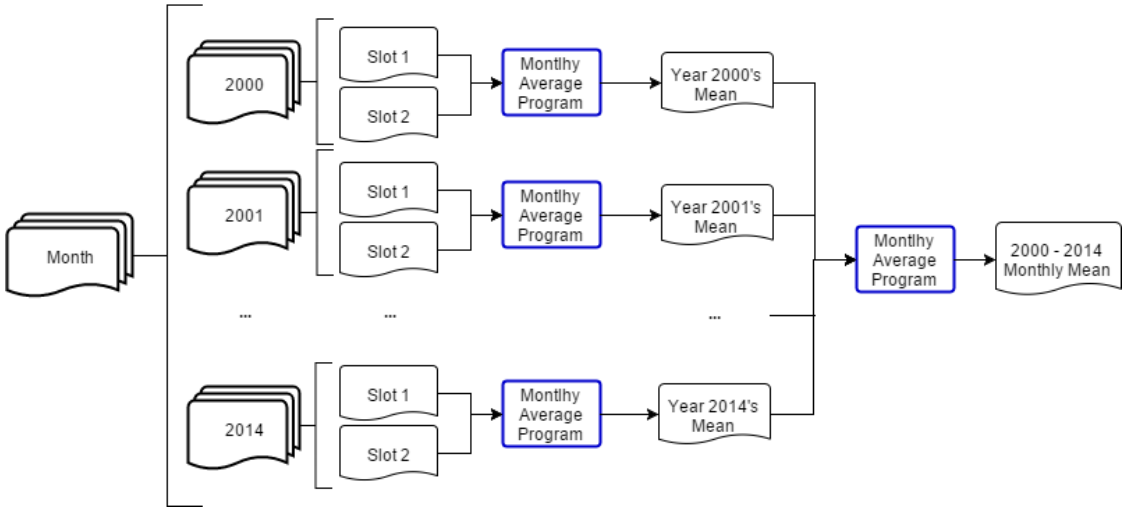


Figure 3.4 | Schematic representation of NDVI data pre-processing for each month.

3.3.1.2. Land Surface Temperature

The MOD11A2 product is distributed in adjacent non-overlapping tiles that make up a world sinusoidal tiled grid with columns and lines. Bayankhongor province is placed between two tiles, line 4 columns 24 and 25 (v04h24 and v04h25), as highlighted in Figure 3.5, with yellow squares.

[6] WGS 84 / UTM zone 47N, situated between 96°E and 102°E, northern hemisphere and between equator and 84°N.

This means that LST data had to be downloaded for the two tiles and later put together as a mosaic.

First, the subdata sets had to be extracted from the MOD11A2 HDF files, re-projected to the EPSG 32647 coordinate systems and converted to GeoTIFF format. To do this operation the Python program in Convert HDF to TIF (LST) was developed by using QGIS's extension function *gdalwarp*. This program also mosaics the output v04h24 and v04h25 tiles after extracting and re-projecting the subdata sets. The following step included the clipping of the study area (Appendix A.4) from the raster files generated based on a shapefile from Bayankhongor province made available by Global Administrative Areas (GADM) on the link http://www.gadm.org/country.

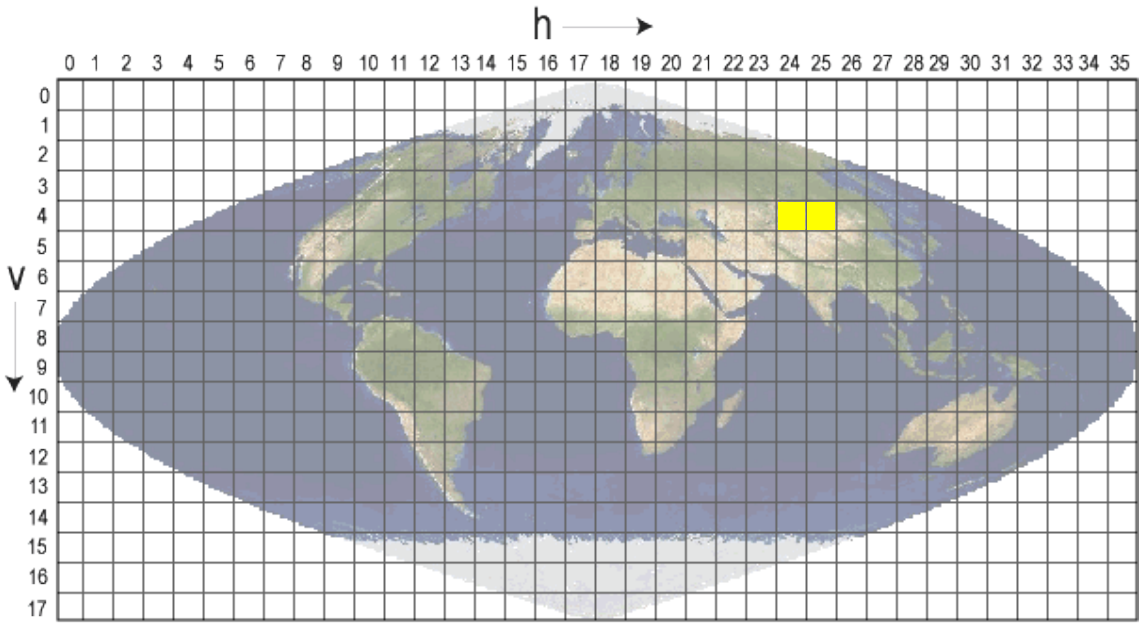


Figure 3.5 | Sinusoidal world tile grid with Bayankhongor tiles in yellow. (National Aeronautics and Space Administration, Goddard Space Flight Center, 2013)

The digital numbers of each pixel in the LST layers needed to be converted to temperature units (Kelvin) by multiplying by a scale factor of 0,02 (Land Processes Distributed Active Archive Center (LP DAAC), 2014). In order to convert Kelvin to Celsius, it was used the Equation 3 -2 (National Institute of Standards and Technology, 2008):

$$Celcius = Kelvin - 273,15 \qquad \textbf{Equation 3-2}$$

The LST layers were then filtered based on the QC layers. Accordingly to (Fondazione Edmund Mach (FEM), 2015) the pixel values that represent a lower quality correspond to integers 2, 3, equal and greater than 129. These values were masked by assigning them a no data value (-9999). Python program in Appendix A.5 was used to apply the QC mask and convert the temperature units. Similar to NDVI data, LST 8-day images were also divided into monthly folders in order to calculate the monthly mean for each year from 2000 to 2014 and the overall mean for the long term period.

3.3.1.3. Precipitation

Precipitation monthly data from TRMM3B43 was obtained in NetCDF format. Since the data was already in monthly format a single Python program was developed to do most of the pre-processing. Thus, the program in the Convert NC to TIF, Clip and Calculate Mean (Precipitation), extracts the precipitation rate layers from the 3B43 NetCDF files, converts them to GeoTIFF format and re-projects them to the EPSG 32647 coordinate system. The same program then clips the Bayankhongor province from the global images.

In order to obtain a more understandable and relevant analysis of precipitation data, it was calculated the monthly accumulated rainfall by using Equation 3-3. The program in Appendix A.5 was used to generate the monthly accumulated precipitation images.

$$\boldsymbol{Accumulated\ rainfall\ = Pr\ x\ 24\ x\ D} \qquad \textbf{Equation 3-3}$$

Where:

- Pr, precipitation rate (hourly rainfall)
- 24, hours in one day
- D, days in each month

3.3.2. Data Analysis

3.3.2.1. Anomalies

Having calculated the monthly composites for NDVI and LST, it was possible to generate anomaly images for every data set for each month during the 2000-2014 period. Anomaly images were calculated based on the formula in Equation 3-4, applied to monthly NDVI, LST and Precipitation images through the program in Anomaly calculation

$$\boldsymbol{Anomaly_{my}\ = DS_{my} - DS_{average\ m2000-2014}} \qquad \textbf{Equation 3-4}$$

Where:

- m, month
- y, year
- DS, data set.

3.3.2.2. Correlation

All datasets being analysed exhibit different spatial resolutions. In order to correctly have any confidence in the correlation results, the coarser resolution was the highest level of detail being studied. Thus, NDVI and LST anomalies were resampled to Precipitation's anomalies resolution. After going through Precipitation images metadata it was concluded that the resolution was approximately 16 km (15935,5 km) which is a better resolution than the initial 27 km. The resampled anomaly images were generated by the program in Resample to coarser resolution

Pearson's correlation coefficient calculations were conducted between NDVI and LST and NDVI and Precipitation. Each month was analysed separately as a result of relationship variations between NDVI and the other data during the growing season (Ji & Peters, 2003). The correlations were calculated for three known drought years in Mongolia: 2001, 2002 and 2009 (John, et al., 2013). One of Pearson's correlation coefficient formulas is:

$$Coef = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(x_i - \bar{x})^2}} \qquad \textbf{Equation 3-5}$$

The results were obtained by means of the program in Appendix A.8 which applies the *pearsonr()* function. This function calculates the coefficient, the p-value used for testing non-correlation between two raster images with the same resolution.

3.3.2.3. Growing season temporal correlations for drought years

Temporal correlations for 2000, 2002 and 2009's growing seasons were obtained through Appendix A.9's program that makes pixel by pixel correlations. The program allows access to the layer stack of just one pixel for one dataset and compare it with the corresponding pixel on the other layer stack, has shown in Figure 3.6.

The correlations calculated in the program are also based in the equation 3-5 and use the function *pearsonr().* The output generates three images: (1) temporal correlation, (2) p-value and (3) significant values, which correspond to p-values under 0,1.

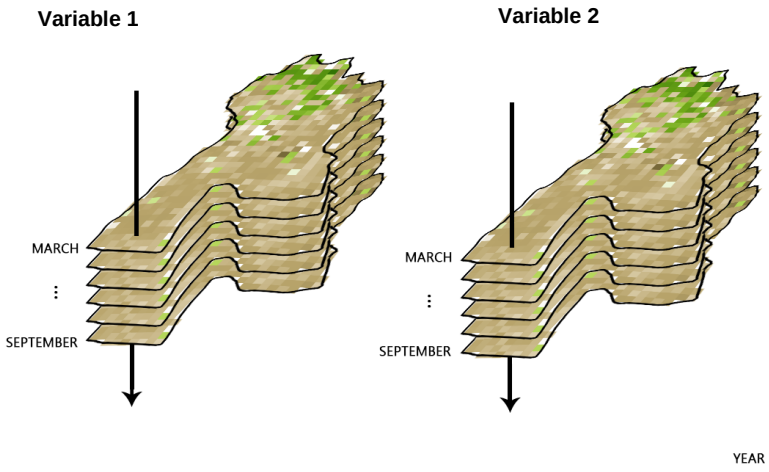


Figure 3.6 | Schematic representation of the concept behind the temporal correlation between two datasets.

4. Results and Discussion

4.1. Exploratory Analysis

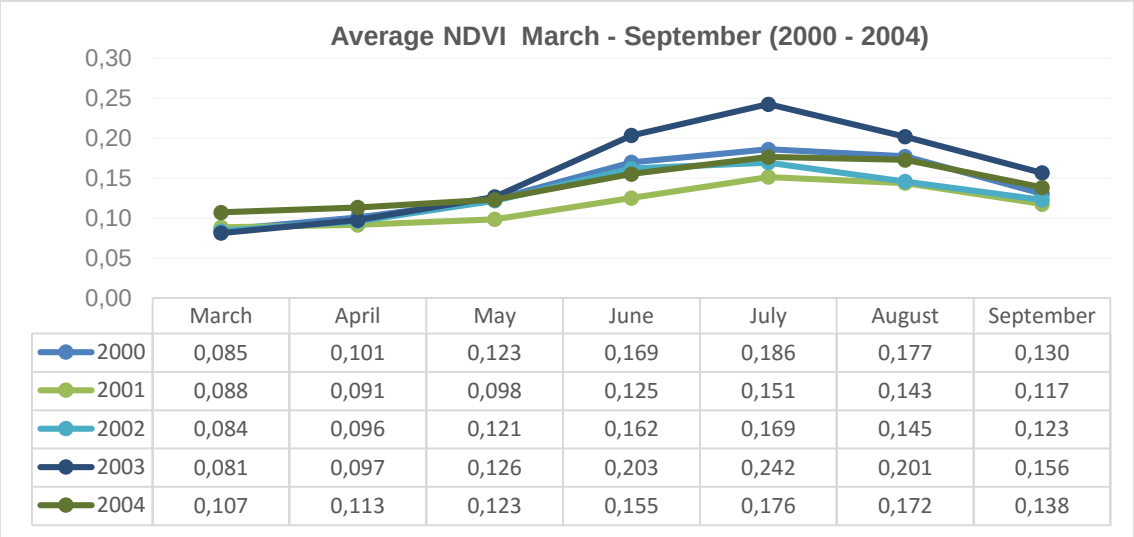
4.1.1. NDVI

As an overall analysis, the annual pattern of NDVI follows a trend according to what's expected for a region in the northern hemisphere. The beginning of the growing season shows the lowest values of NDVI and the highest values are reached during the summer months (June, July and August). However, the range of values (0, 0 to 0,20) throughout the growing season for the fifteen year period is remarkably low (Figure 4.1), mainly due to the type of ecosystems present in Bayankhongor province. Since the territory is covered by desert-like areas with sparse vegetation (Figure 3.1), the photosynthetic activity from plants is very low when compared to other regions. Further in the study this may be reflected in low correlation results between NDVI and the other variables.

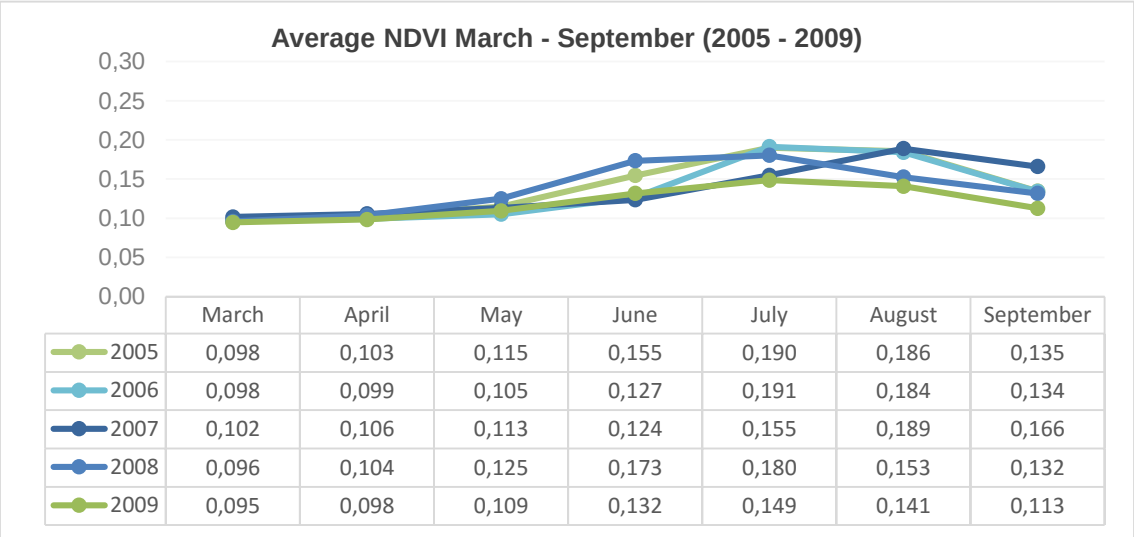
The year of 2003 displays the highest seasonal amplitude of the time series with a peak value of 0,242 in July. The lowest values can be observed during 2001 and 2009 where the minimum value is 0,088 and the maximum value equals to 0,151. From 2005 to 2014, excluding 2009, NDVI appears to be increasing in Bayankhongor province which translates into a growth in vegetation's photosynthetic activity.

The results obtained for NDVI monthly anomalies for the period 2000 – 2014 are plotted in Appendix B. Four months throughout the growing season (March, May, July and September) were selected for analysis (Figure 4.2 | NDVI Anomalies for March, May, July and SeptemberFigure 4.2). Also, the spatial distribution of NDVI anomalies in Bayankhongor province is represented in Figure 4.3 and Figure 4.4**Erro! A origem da referência não foi encontrada.** for the same four months, for a standard year (2006), a dry year (2001) and a wet year (2003). These years were chosen based on the average values of the growing season for each year.

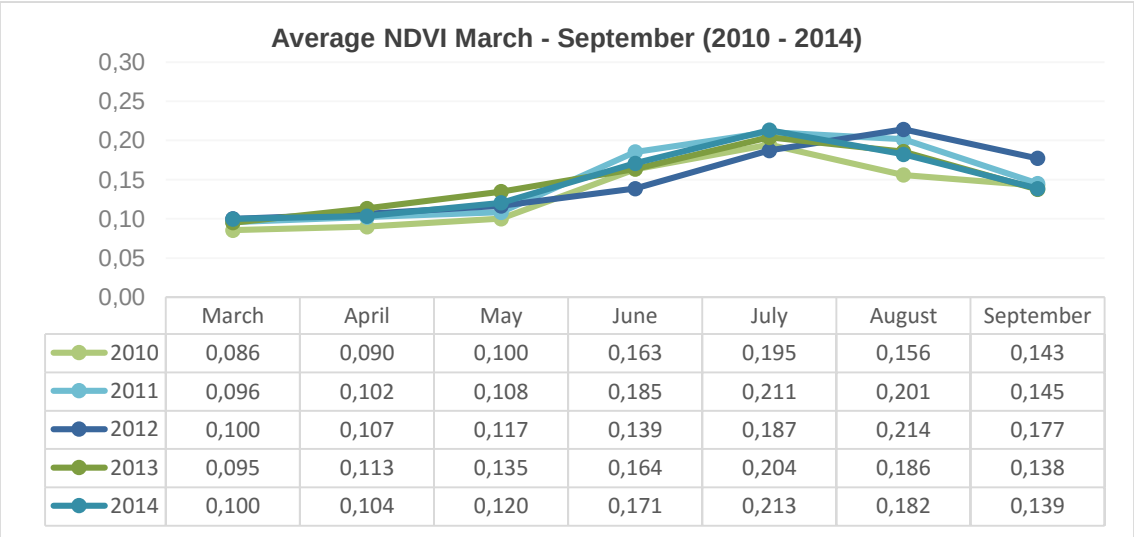
The year of 2001 exhibits negative anomalies all the way through the growing season, being July the month with the highest negative anomaly. On the contrary, the year of 2003 is characterized by positive anomalies, indicating a vegetation growth in comparison to the other years (Figure 4.2). This can be confirmed by the anomaly images in Figure 4.3 and Figure 4.4. Both 2001 and 2003 images are completely different from the standard year for all the selected months. In 2001, the vegetation in the majority of Bayankhongor province's territory decreased, except in some northern areas where the Khangai Mountain Ranges lay, which are characterized by cooler temperatures and higher amount of precipitation. The year of 2003 registered the highest positive anomalies, which cover the entire territory with the exception of some areas in the south region that refer to the desert biome accordingly to Figure 3.1. Based solely on the anomaly images, NDVI represents well the drought years during the studied period.



	March	April	May	June	July	August	September
2000	0,085	0,101	0,123	0,169	0,186	0,177	0,130
2001	0,088	0,091	0,098	0,125	0,151	0,143	0,117
2002	0,084	0,096	0,121	0,162	0,169	0,145	0,123
2003	0,081	0,097	0,126	0,203	0,242	0,201	0,156
2004	0,107	0,113	0,123	0,155	0,176	0,172	0,138



	March	April	May	June	July	August	September
2005	0,098	0,103	0,115	0,155	0,190	0,186	0,135
2006	0,098	0,099	0,105	0,127	0,191	0,184	0,134
2007	0,102	0,106	0,113	0,124	0,155	0,189	0,166
2008	0,096	0,104	0,125	0,173	0,180	0,153	0,132
2009	0,095	0,098	0,109	0,132	0,149	0,141	0,113



	March	April	May	June	July	August	September
2010	0,086	0,090	0,100	0,163	0,195	0,156	0,143
2011	0,096	0,102	0,108	0,185	0,211	0,201	0,145
2012	0,100	0,107	0,117	0,139	0,187	0,214	0,177
2013	0,095	0,113	0,135	0,164	0,204	0,186	0,138
2014	0,100	0,104	0,120	0,171	0,213	0,182	0,139

Figure 4.1 | Normalized Difference Vegetation Index monthly averages for the time period 2000-2014.

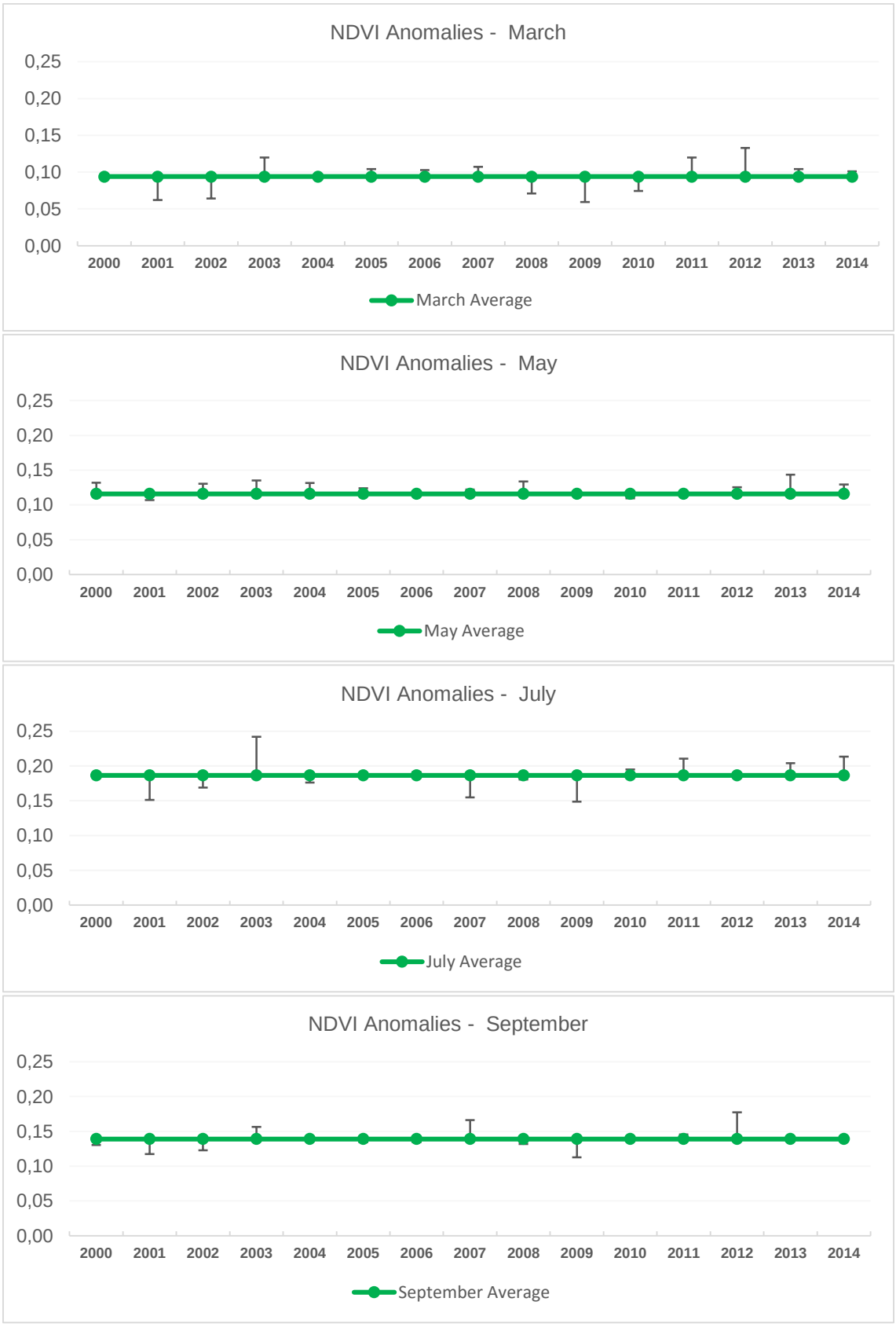


Figure 4.2 | NDVI Anomalies for March, May, July and September, with reference to the respective monthly average for the period 2000-2014.

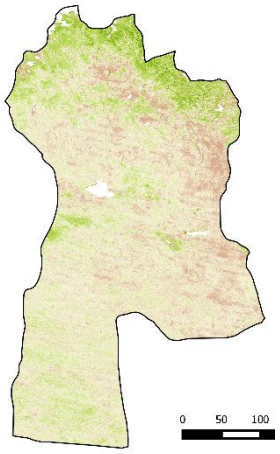
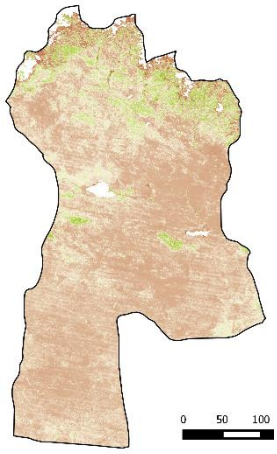
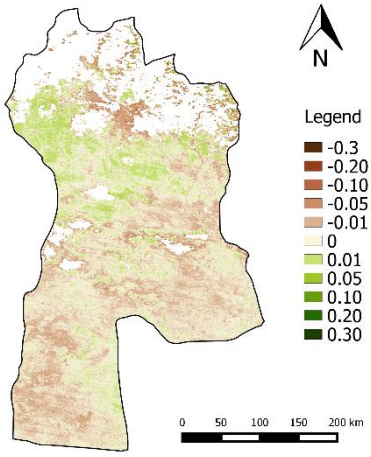
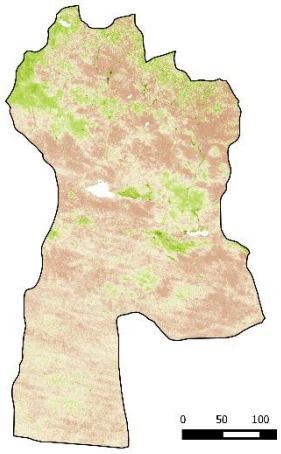
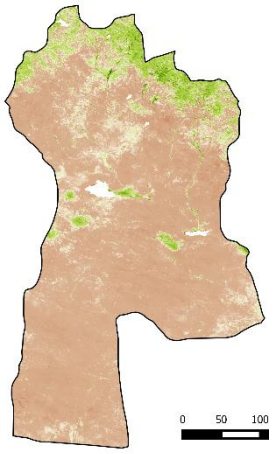
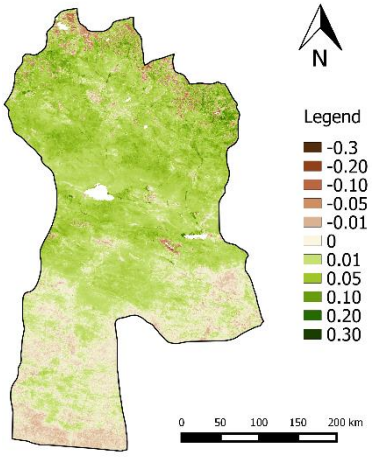
Month	Standard Year (2006)	Dry Year (2001)	Wet Year (2003)
March			
May			

Figure 4.3 | Spatial distribution of NDVI anomalies for a standard, dry and wet year (March and May)

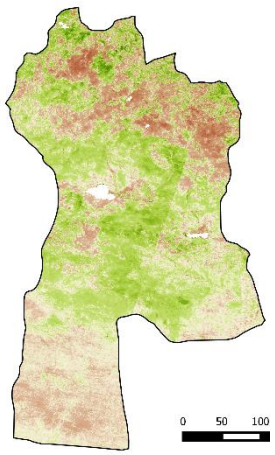
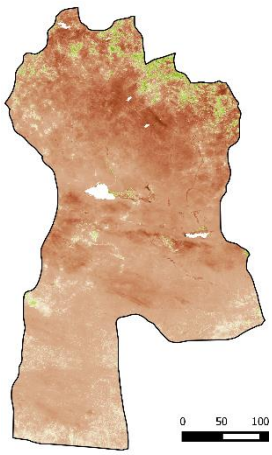
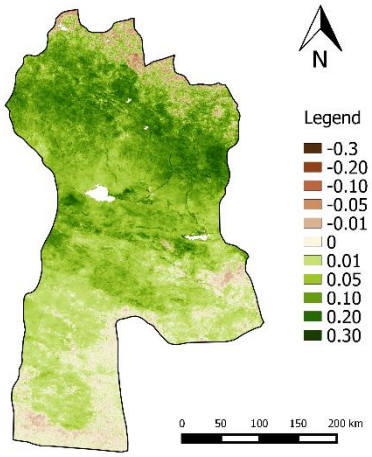
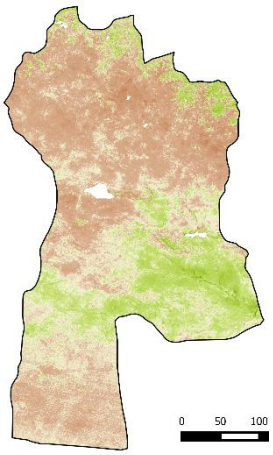
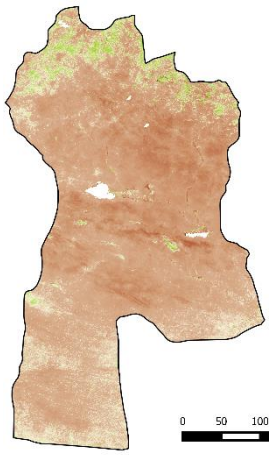
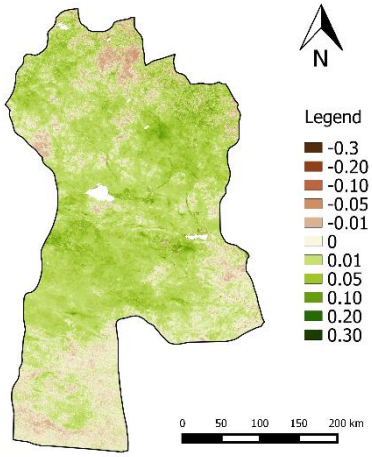
Month	Average Year (2006)	Dry Year (2001)	Wet Year (2003)
July			
September			

Figure 4.4 | Spatial distribution of NDVI anomalies for a standard, dry and wet year (July and September)

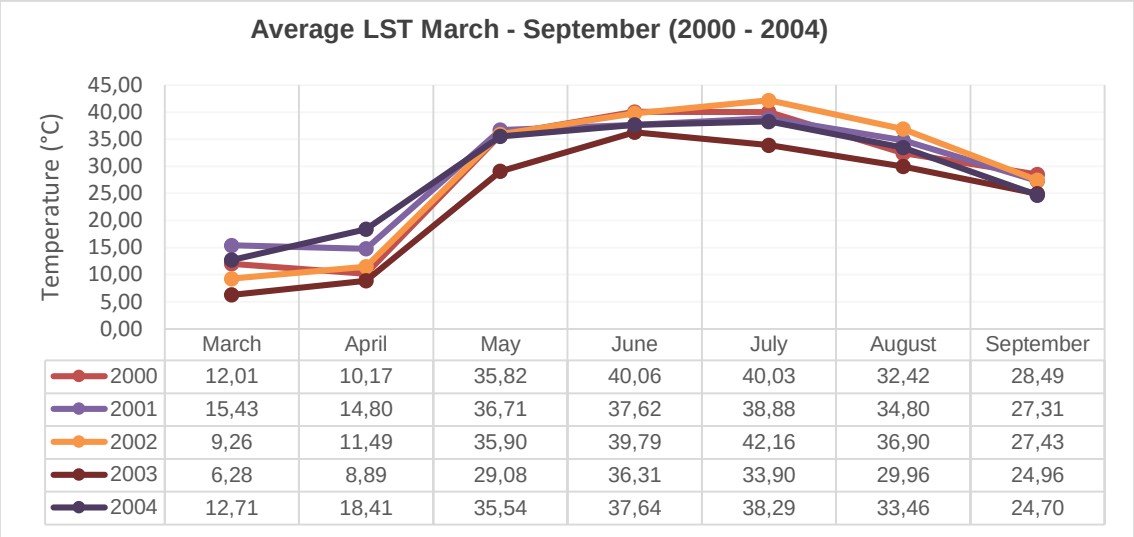
4.1.2. Land Surface Temperature

Average Land Surface Temperature results can be seen in Figure 4.5. Along the growing season LST values rise from March to June or July and start decreasing until September. Thus, the highest temperatures are experienced in May, June, July and August. LST is different from air temperature (temperature measured usually 1,5 m above ground) as it refers to a measurement of land temperature which heats and cools more rapidly than air. This explains the high values seen in the graphic plots as the growing season covers the summer month.

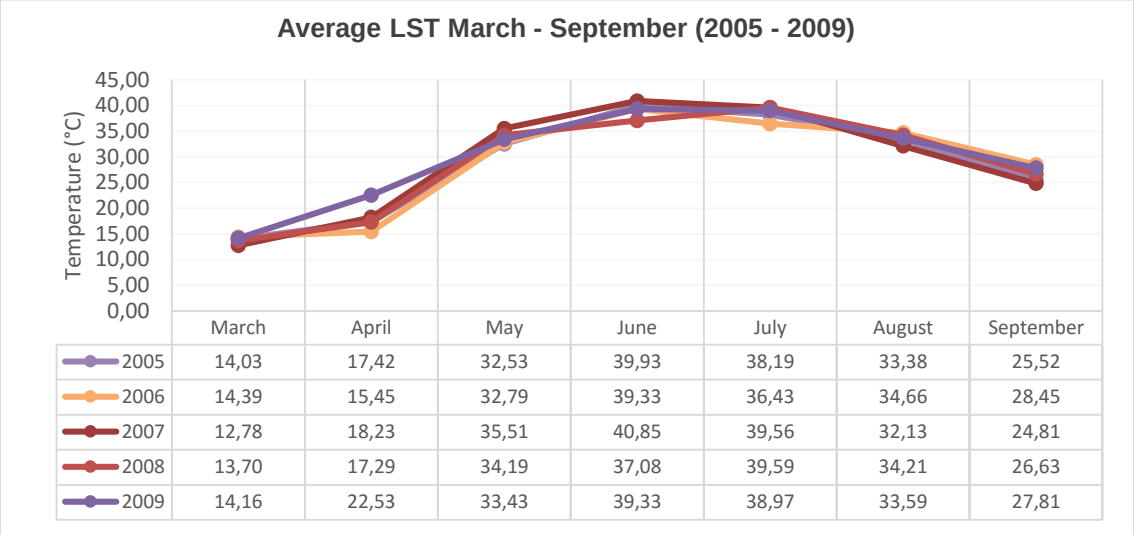
Although LST values show similar trends for along the years it is possible to identify warmer and cooler years from the time series. In the years between 2000 and 2002, LST reached values of 40°C, which can be explained by the summer droughts experienced in Mongolia during that time. LST values drop in 2003, making it the coolest year of the fifteen year period. The years of 2007 and 2009 are both characterized by warm temperatures, particularly 2009 in the beginning and end of the growing season. From 2000 to 2014, LST values have decreased, ranging from 31,4°C to 39,5°C during the summer months (June, July and August).

The average monthly anomalies for each year are plotted in the Appendix D while the spatial distribution of the same anomalies are presented in Appendix E. Overall, the years showing more positive anomalies during the growing season, meaning that the temperature is above average, are 2001, 2002, 2007 and 2009. The year of 2002 has the highest positive anomalies for June, July and August although is shows negative anomalies for March and April. The comparison between a standard year, a dry year and a wet year for four chosen months can be seen in Figure 4.6 and Figure 4.7.

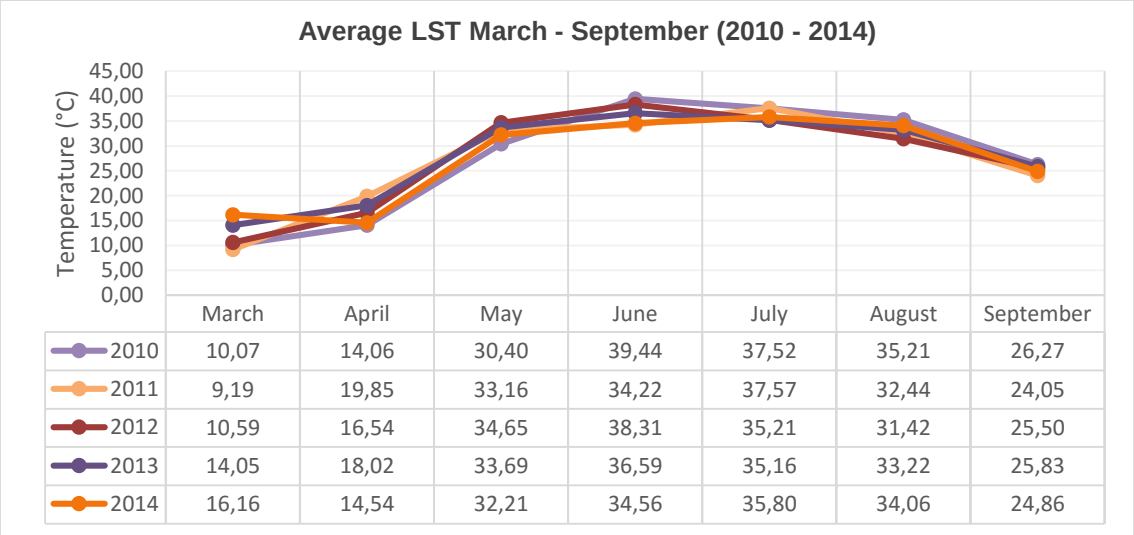
At the beginning of the growing season (March) it is noticeable an increase of the LST in the standard and dry years (Figure 4.6). The year of 2001 (dry year) LST increase is distributed over all territory of Bayankhongor province, while in 2006 the desert biome shows negative anomalies. This corroborates the average values plotted in Figure 4.5 as 2001 has the highest value for March. Throughout the growing season the wet year (2003) stands out when compared to the other due to the high negative anomalies, especially in the steppe and mountainous region (Figure 3.1). It is also worth noting that at the end of the growing season (September) both standard and dry years exhibit positive anomalies, although they’re not as significant as in March (Figure 4.7).



	March	April	May	June	July	August	September
2000	12,01	10,17	35,82	40,06	40,03	32,42	28,49
2001	15,43	14,80	36,71	37,62	38,88	34,80	27,31
2002	9,26	11,49	35,90	39,79	42,16	36,90	27,43
2003	6,28	8,89	29,08	36,31	33,90	29,96	24,96
2004	12,71	18,41	35,54	37,64	38,29	33,46	24,70



	March	April	May	June	July	August	September
2005	14,03	17,42	32,53	39,93	38,19	33,38	25,52
2006	14,39	15,45	32,79	39,33	36,43	34,66	28,45
2007	12,78	18,23	35,51	40,85	39,56	32,13	24,81
2008	13,70	17,29	34,19	37,08	39,59	34,21	26,63
2009	14,16	22,53	33,43	39,33	38,97	33,59	27,81



	March	April	May	June	July	August	September
2010	10,07	14,06	30,40	39,44	37,52	35,21	26,27
2011	9,19	19,85	33,16	34,22	37,57	32,44	24,05
2012	10,59	16,54	34,65	38,31	35,21	31,42	25,50
2013	14,05	18,02	33,69	36,59	35,16	33,22	25,83
2014	16,16	14,54	32,21	34,56	35,80	34,06	24,86

Figure 4.5 | Land Surface Temperature monthly averages for the time period 2000 – 2014.

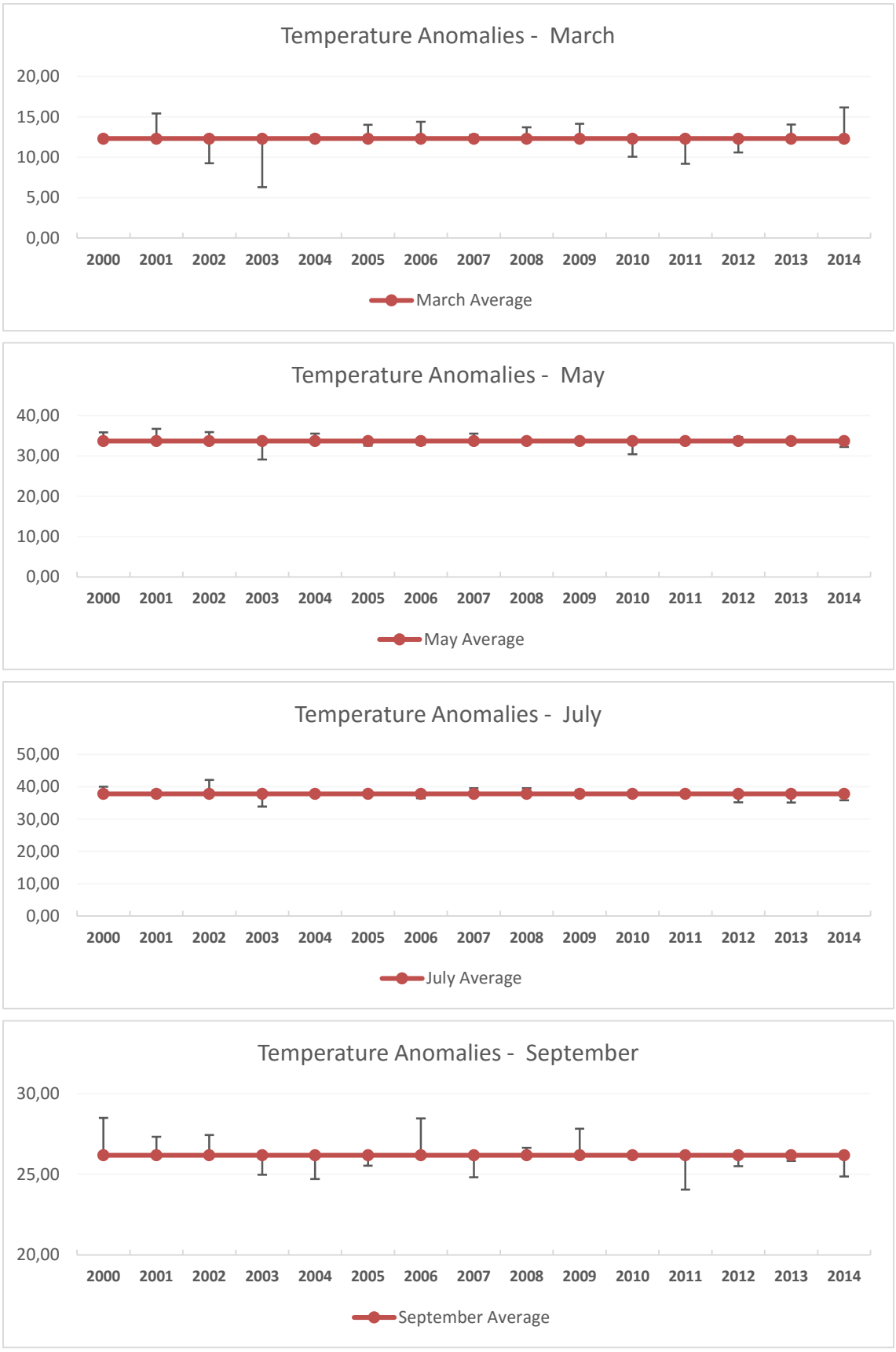


Figure 4.6 | LST Anomalies (°C) for March, May, July and September, with reference to the respective monthly average for the period 2000-2014.

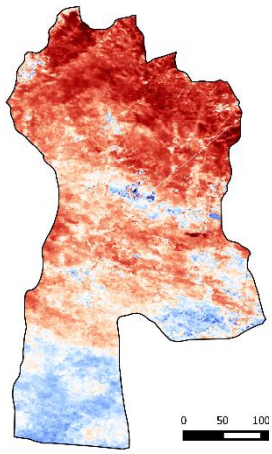
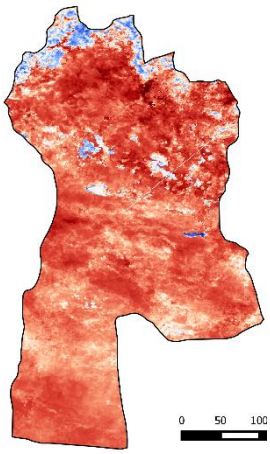
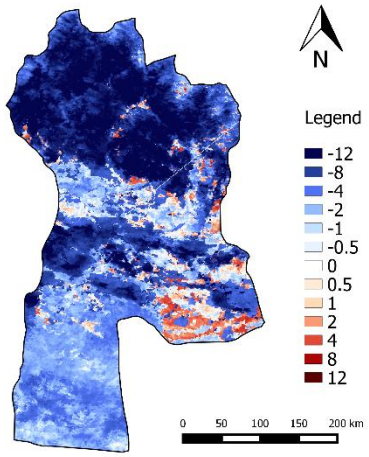
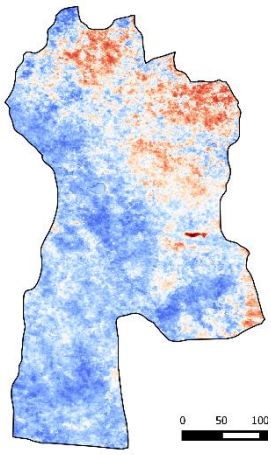
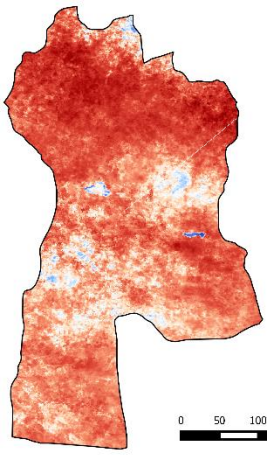
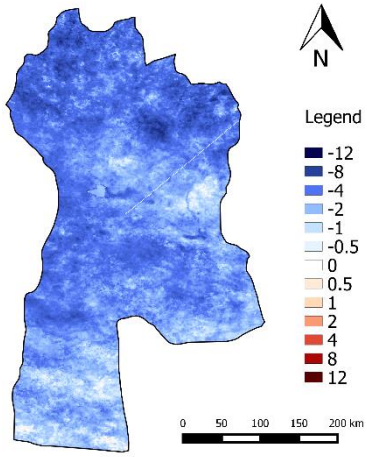
Month	Standard Year (2006)	Dry Year (2001)	Wet Year (2003)
March			
May			

Figure 4.7 | Spatial distribution of LST anomalies (°C) for a standard, dry and wet year, for the months of March and May.

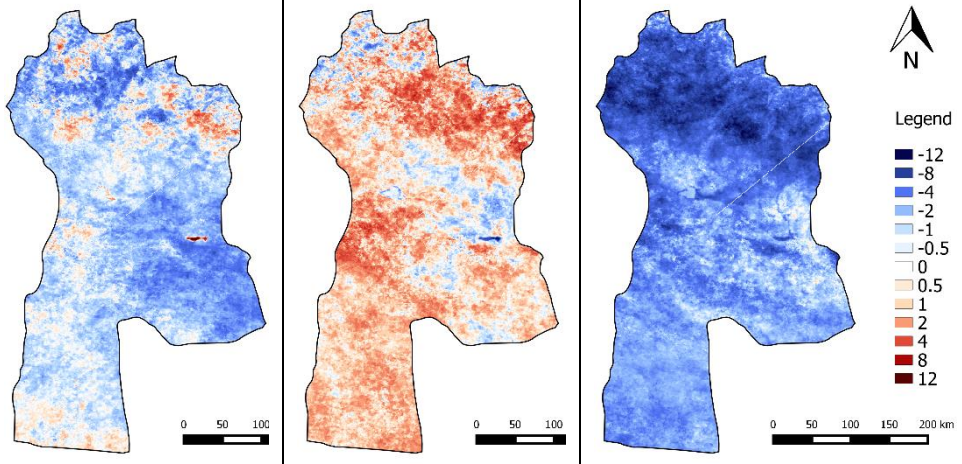
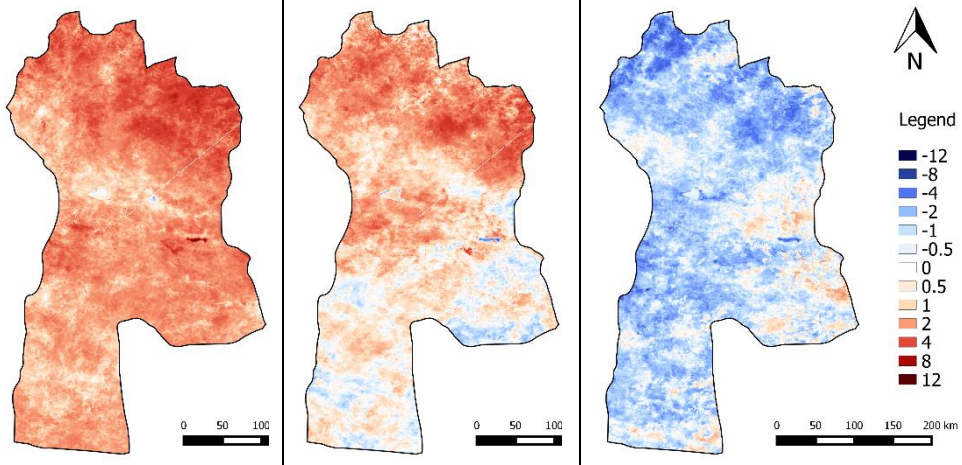
Month	Standard Year (2006)	Dry Year (2001)	Wet Year (2003)
July			
September			

Figure 4.8 | Spatial distribution of LST anomalies (°C) for a standard, dry and wet year, for the months of July and September.

4.1.3. Precipitation

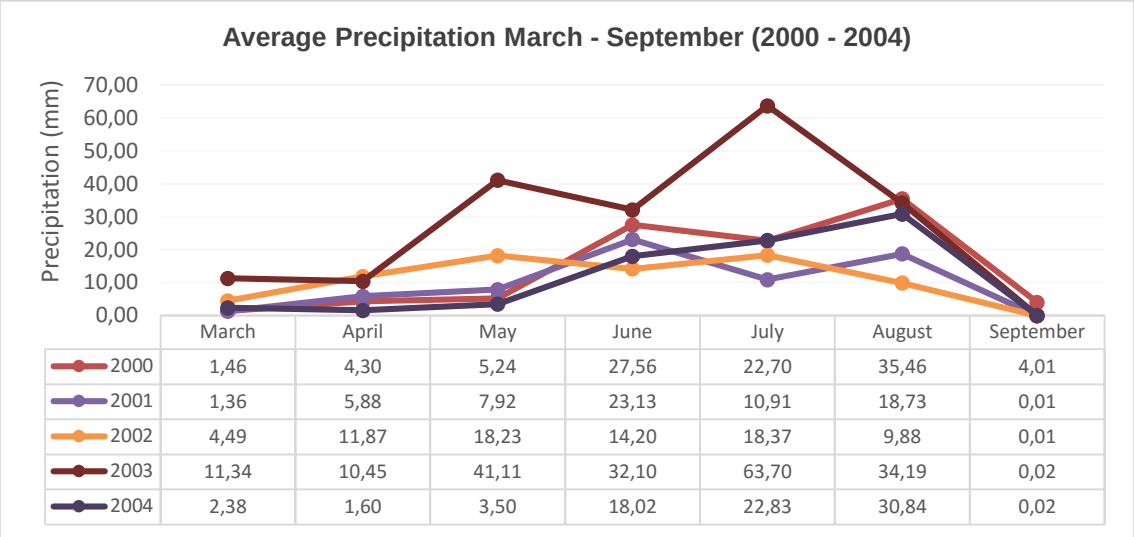
The monthly average of precipitation during the growing season is represented in Figure 4.9. Although precipitation's variability through the years, it still follows a seasonal pattern in accordance Mongolia's region's rainy season. In general, from March to May the precipitation is very low in contrast with June, July and August when it registers the highest values.

From 2000 to 2002, Mongolia was going through a summer-*dzud* which can explain the lower precipitation values observed in these years. The year 2003 shows atypical values when compared to others since it registers the highest peaks of the time series (May and July). While 2005, 2006 and 2007 also show high values of precipitation during the summer months, in 2008 and 2009 these values decrease, being 2009 the driest year of this group. The precipitation increases from 2010 to 2014, having three years (2011, 2012 and 2014) showing values above 50 mm during the summer.

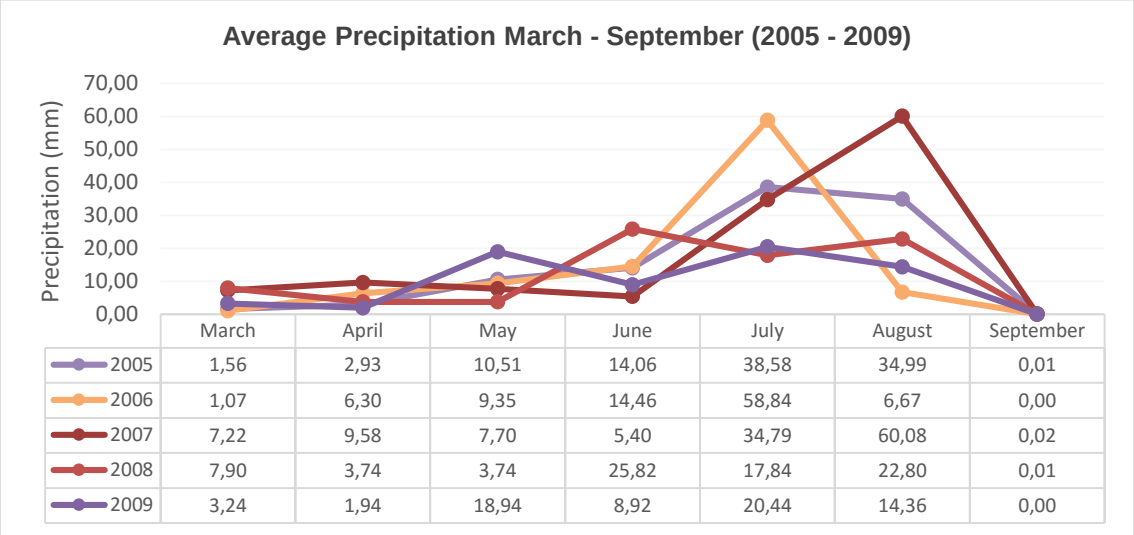
In general, the precipitation values are extremely low but expected, taking into account the land cover in Bayankhongor province. The majority of the territory is covered by desert and desert steppe (Figure 3.1), where the annual precipitation ranges from 0 to 100 mm. Thus, the results can be considered representative of the study area. The precipitation anomalies for the time-series period are represented in Figure 4.9 and their spatial distribution for March, May, July and September are exhibited in Figure 4.10 and Figure 4.11.

The wet year (2003) displays an increase of precipitation throughout the growing season, with highest values in May and July as expected. The positive anomalies cover the majority of the territory, except for the month of September which is characterized by a decrease of precipitation in the north and an increase of precipitation in the central area (desert steppe). The average values of these anomalies, plotted in Figure 4.9, corroborate the images shown in Figure 4.10 and Figure 4.11 since that for the year of 2003 the anomalies are positive with exception of September.

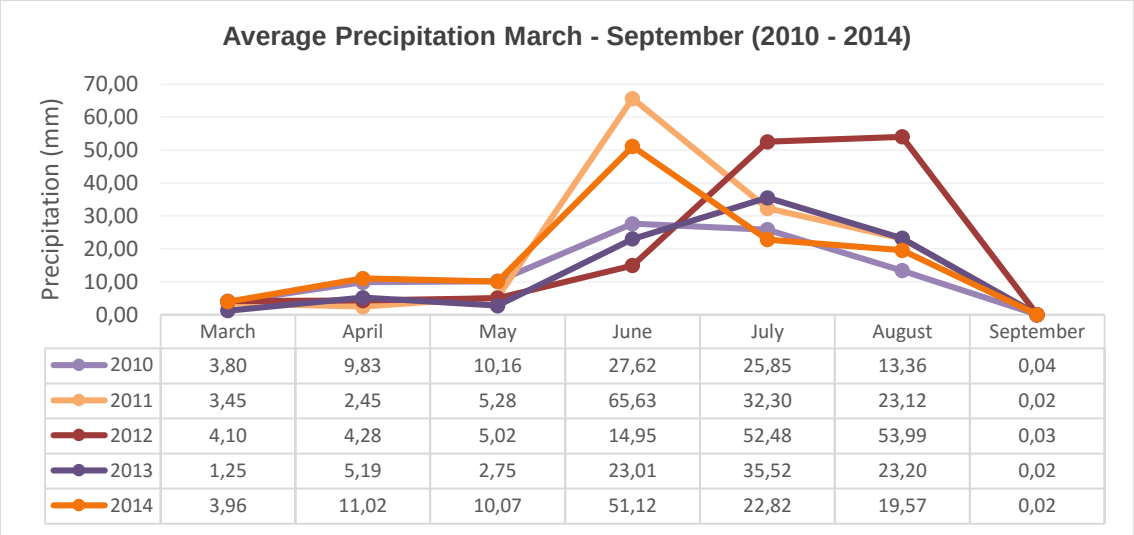
The standard year (2006) reveals a decrease in precipitation values at the beginning of the growing season and an abrupt increase of precipitation in July. This can be also noticed in Figure 4.9, where for the month of July, the year of 2006 registers a high positive anomaly. On the contrary, the year of 2001 (dry year) registers a negative anomaly which translates into the image in Figure 4.10 where the decrease of precipitation is distributed over the whole province.



	March	April	May	June	July	August	September
2000	1,46	4,30	5,24	27,56	22,70	35,46	4,01
2001	1,36	5,88	7,92	23,13	10,91	18,73	0,01
2002	4,49	11,87	18,23	14,20	18,37	9,88	0,01
2003	11,34	10,45	41,11	32,10	63,70	34,19	0,02
2004	2,38	1,60	3,50	18,02	22,83	30,84	0,02



	March	April	May	June	July	August	September
2005	1,56	2,93	10,51	14,06	38,58	34,99	0,01
2006	1,07	6,30	9,35	14,46	58,84	6,67	0,00
2007	7,22	9,58	7,70	5,40	34,79	60,08	0,02
2008	7,90	3,74	3,74	25,82	17,84	22,80	0,01
2009	3,24	1,94	18,94	8,92	20,44	14,36	0,00



	March	April	May	June	July	August	September
2010	3,80	9,83	10,16	27,62	25,85	13,36	0,04
2011	3,45	2,45	5,28	65,63	32,30	23,12	0,02
2012	4,10	4,28	5,02	14,95	52,48	53,99	0,03
2013	1,25	5,19	2,75	23,01	35,52	23,20	0,02
2014	3,96	11,02	10,07	51,12	22,82	19,57	0,02

Figure 4.9 | Precipitation monthly averages for the time period 2000 – 2014.

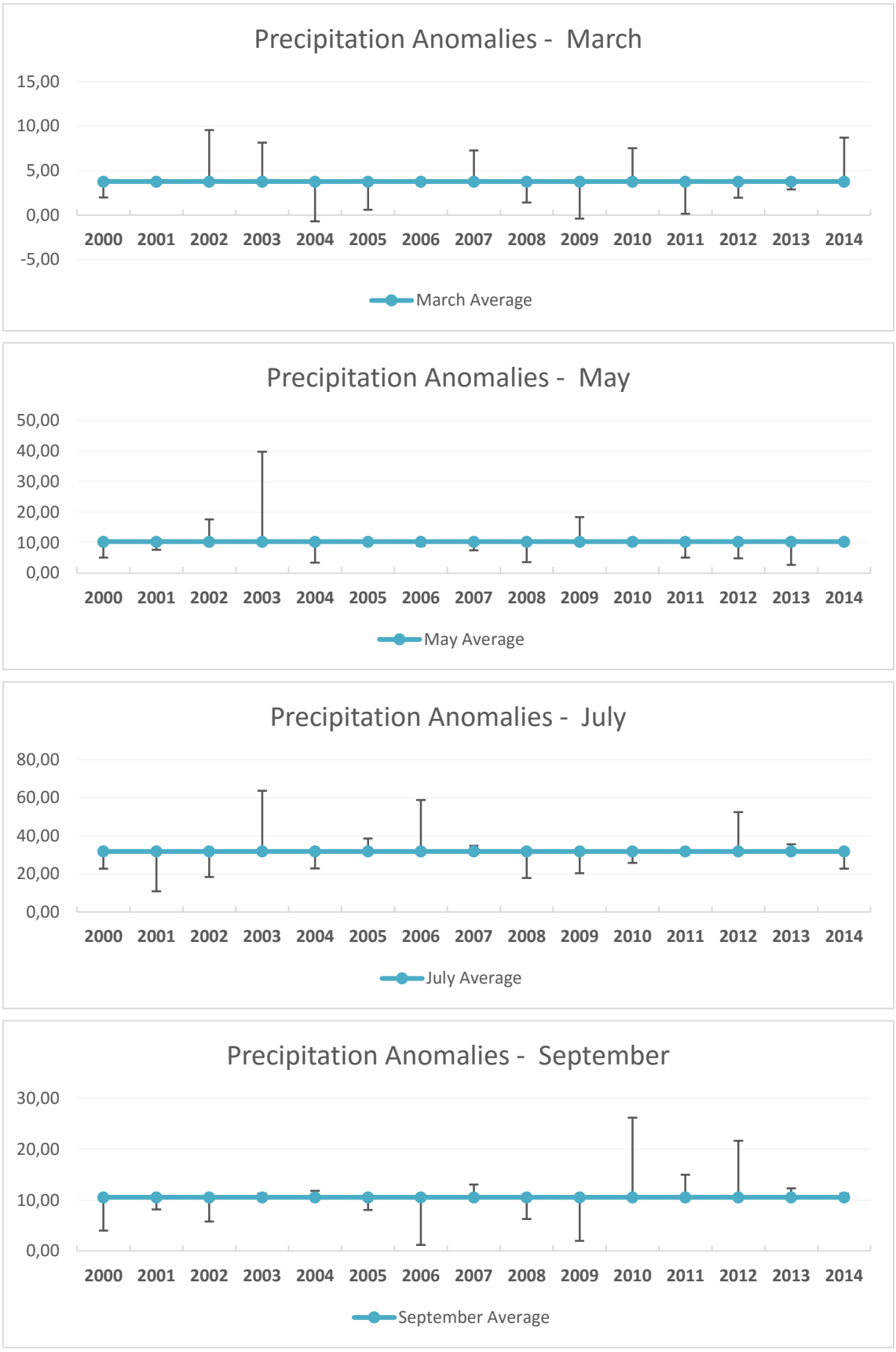


Figure 4.10 | Precipitation Anomalies (mm) for March, May, July and September, with reference to the respective monthly average for the period 2000-2014.

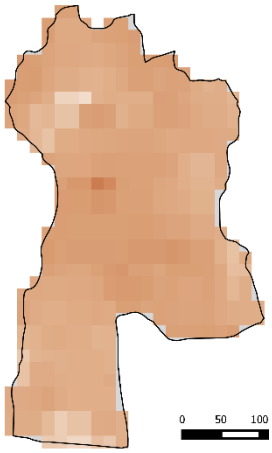
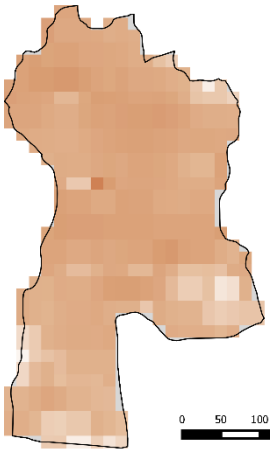
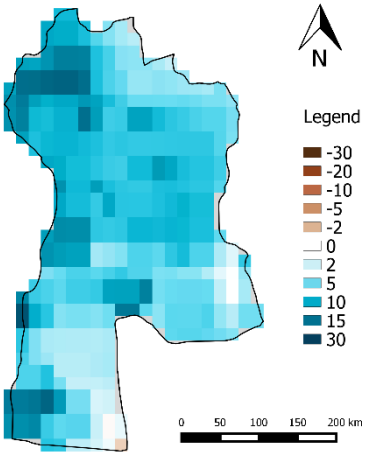
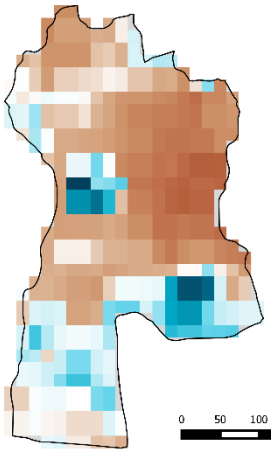
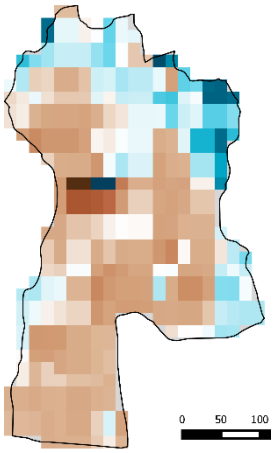
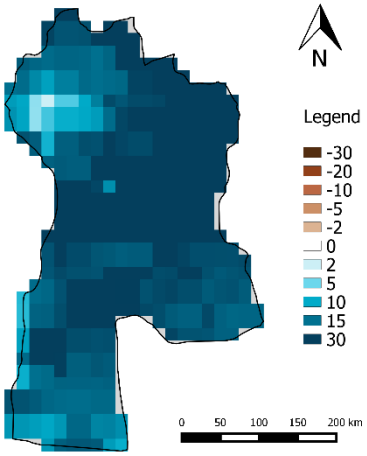
Month	Standard Year (2006)	Dry Year (2001)	Wet Year (2003)
March			
May			

Figure 4.11 | Spatial distribution of precipitation anomalies (mm) for a standard, dry and wet year, for the months of March and May.

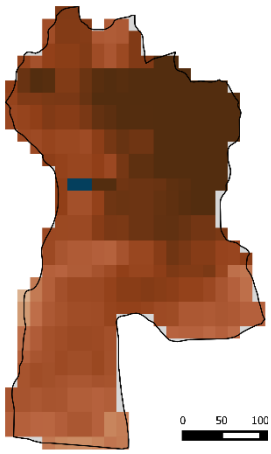
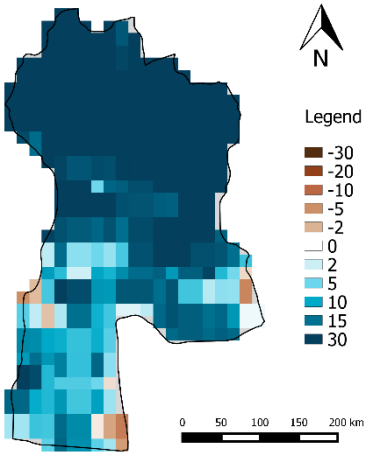
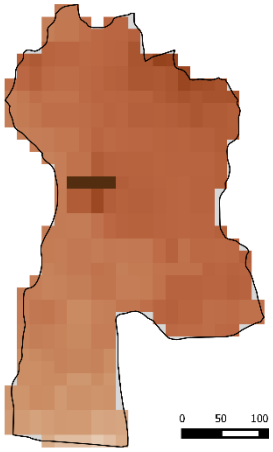
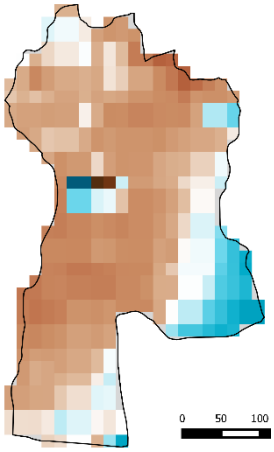
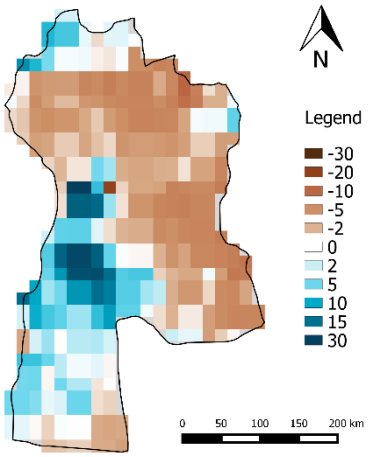
Month	Standard Year (2006)	Dry Year (2001)	Wet Year (2003)
July			
September			

Figure 4.12 | Spatial distribution of precipitation anomalies (mm) for a standard, dry and wet year, for the months of July and September.

4.2. Monthly correlations of data sets anomalies for drought years

The results of monthly correlations between NDVI and LST and NDVI and Precipitation are systematized in Table 4-1,

Table 4-2 and Table 4-3 for drought years 2001, 2002 and 2009 respectively. NDVI-LST results have shown mostly negative correlations which means that when the LST values are higher, NDVI values decrease. Contrarily, the NDVI-Precipitation correlations are mostly positive, meaning that NDVI increases with precipitation values as expected. The months showing highest correlation values for both NDVI-LST and NDVI-Precipitation are June, July and August. These are the growing season months with higher NDVI values, as previously shown in Figure 4.1.

It is important to notice that in June the correlations show uncharacteristic results in comparison to the other months. While the NDVI-LST correlations are positive, the NDVI-Precipitation correlations are negative. However this can be explained by the higher values of accumulated precipitation in April and May for these years (Figure 4.9). The precipitation in the previous months may have created conditions to maintain a certain amount of vegetation's photosynthetic activity, even though the LST values were high and the precipitation was limited in June.

Overall the monthly results show weak correlations between the variables, with a few exceptions. Taking this into consideration it is expected growing season temporal images with low correlations and insufficient significant values.

Table 4-1 | Monthly correlation of NDVI, LST and Precipitation anomalies (2001). The significant values are in bold while the highest significant correlations are in red.

	NDVI (LST)		NDVI (Precipitation)	
	Correlation	**P-Value**	**Correlation**	**P-Value**
March	**0,18**	0,00	**-0,09**	0,06
April	-0,03	0,46	0,01	0,78
May	**0,11**	0,02	0,03	0,54
June	0,03	0,54	**0,11**	0,02
July	**-0,18**	0,00	**0,37**	0,00
August	**-0,25**	0,00	**0,22**	0,00
September	**-0,17**	0,00	**0,13**	0,00

Table 4-2 | Monthly correlation of NDVI, LST and Precipitation anomalies (2002). The significant values are in bold while the highest significant correlations are in red.

	NDVI (LST)		NDVI (Precipitation)	
	Correlation	P-Value	Correlation	P-Value
March	-0,01	0,90	0,06	0,20
April	-0,01	0,75	0,00	0,99
May	**-0,22**	0,02	**0,09**	0,06
June	**0,36**	0,00	**-0,47**	0,00
July	**-0,72**	0,00	**0,43**	0,00
August	**-0,67**	0,00	**0,52**	0,00
September	**-0,42**	0,00	**-0,12**	0,01

Table 4-3 | Monthly correlation of NDVI, LST and Precipitation anomalies (2009). The significant values are in bold while the highest significant correlations are in red.

	NDVI (LST)		NDVI (Precipitation)	
	Correlation	P-Value	Correlation	P-Value
March	**-0,14**	0,00	-0,02	0,69
April	**-0,29**	0,00	**0,21**	0,00
May	**-0,11**	0,02	**0,28**	0,00
June	**0,44**	0,00	**-0,19**	0,00
July	**-0,59**	0,00	**0,37**	0,00
August	**-0,32**	0,00	**0,32**	0,00
September	**-0,13**	0,01	**0,38**	0,00

4.3. Spatial Distribution of NDVI correlations

4.3.1. NDVI-LST

The spatial distribution of the correlations between NDVI and LST anomalies are shown in Figure 4.13, Figure 4.14 and Figure 4.15. Globally, the results show a weak correlation between NDVI and LST anomalies. The temporal images generated include the seven months of the growing season. By computing the months with higher correlations (June, July and August) with the others, important information might get lost in the process. This leads to images with lower correlations and fewer significant values. The years of 2002 and 2009 have the highest correlations and significant values, which goes accordingly to the monthly correlation results.

In 2001 and 2002 there are positive correlations in the central and south region (desert steppe and desert) of Bayankhongor province and negative correlations towards the north (steppe and mountains). On the contrary, in 2009 this pattern is inversed revealing positive correlations in the north and negative correlations in the south.

It is important to mention that for these images the significant values follow a p-value $p < 0,1$ instead of $p < 0,05$, in order to provide more information to analyse.

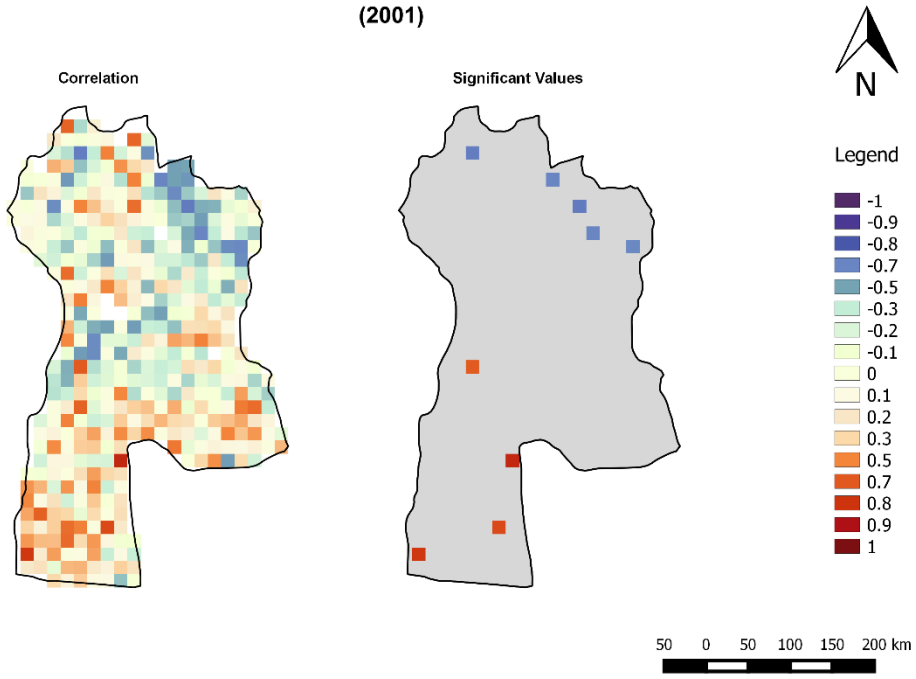


Figure 4.13 | Spatial distribution of temporal correlations between NDVI and LST anomalies for the growing season of 2001. Significant values for $p < 0,1$.

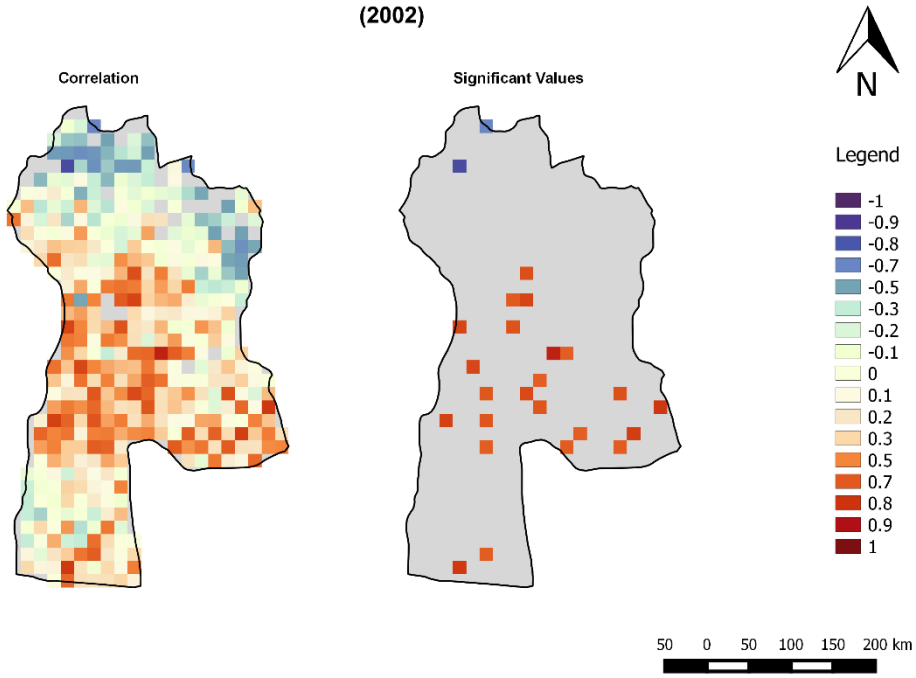


Figure 4.14 | Spatial distribution of temporal correlations between NDVI and LST anomalies for the growing season of 2002. Significant values for p < 0,1.

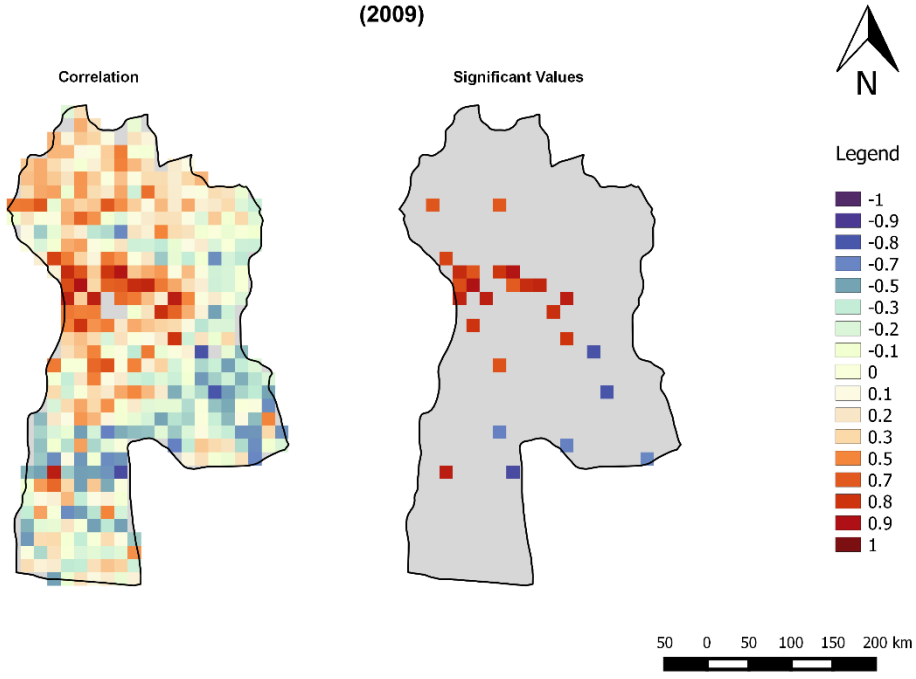


Figure 4.15 | Spatial distribution of temporal correlations between NDVI and LST anomalies for the growing season of 2009. Significant values for p < 0,1.

4.3.2. NDVI-Precipitation

Growing season's spatial distribution of temporal correlations between NDVI and precipitation anomalies are presented in Figure 4.16, Figure 4.17 and Figure 4.18 for 2001, 2002 and 2009 respectively. The results show a high variability between the three drought years, which may lead to inconclusive findings regarding the correlation between NDVI and precipitation anomalies.

The spatial distribution of correlations in 2001 reveals a strong positive correlation between NDVI and precipitation anomalies in the steppe and mountain forest steppe regions. In comparison to the other years, 2001 has more significant values and higher correlations (

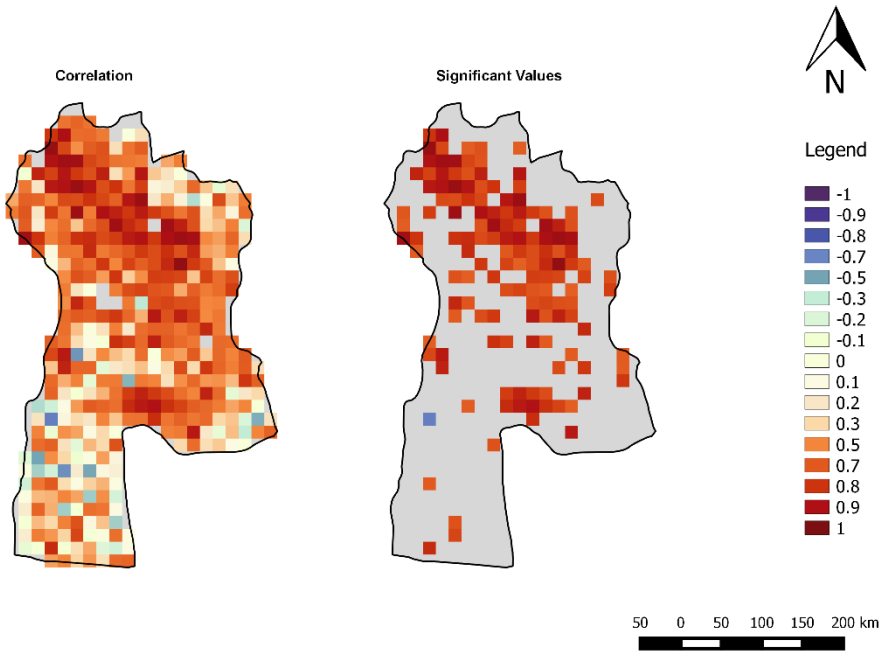
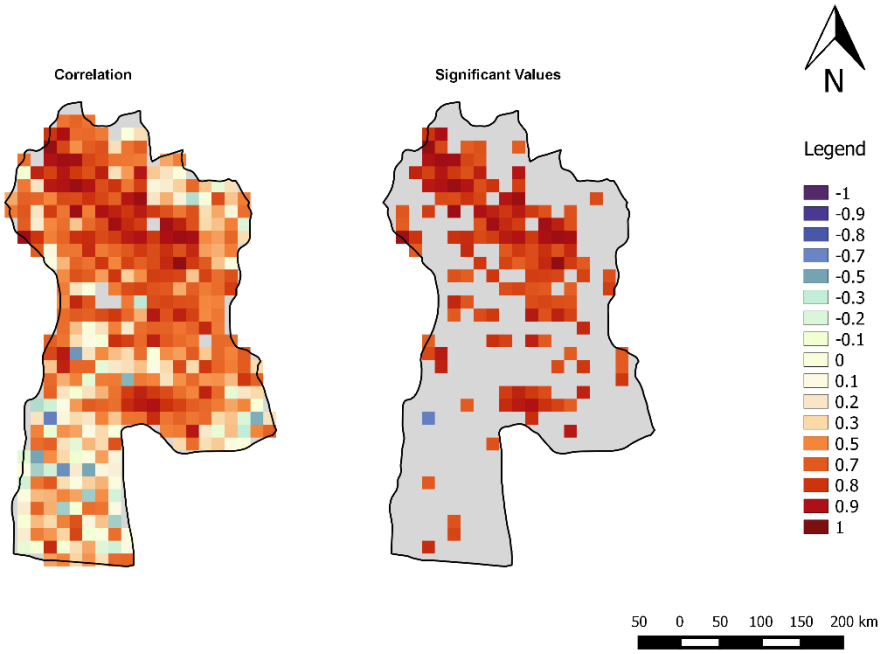


Figure **4.16**). The year of 2009 has the weakest spatial correlation as it has least significant values



(

Figure **4.16**) in comparison with the other two drought years.

Whereas in 2001 and 2009 there are positive correlations in the northern and central region, in 2002 the correlation between NDVI and precipitation anomalies displays negative values towards the south. The year of 2002 is characterized by unusual correlations in June, as seen in

Table 4-2, that are explained by the higher amount of precipitation in the previous months. The images in Figure 4.16, Figure 4.17, and Figure 4.18 were generated by combining the seven months of the growing season. This gathering of information may affect the correlation results since it assumes that all the precipitation in the growing season is going to impact NDVI's response in the same period without considering a time lag between the two variables.

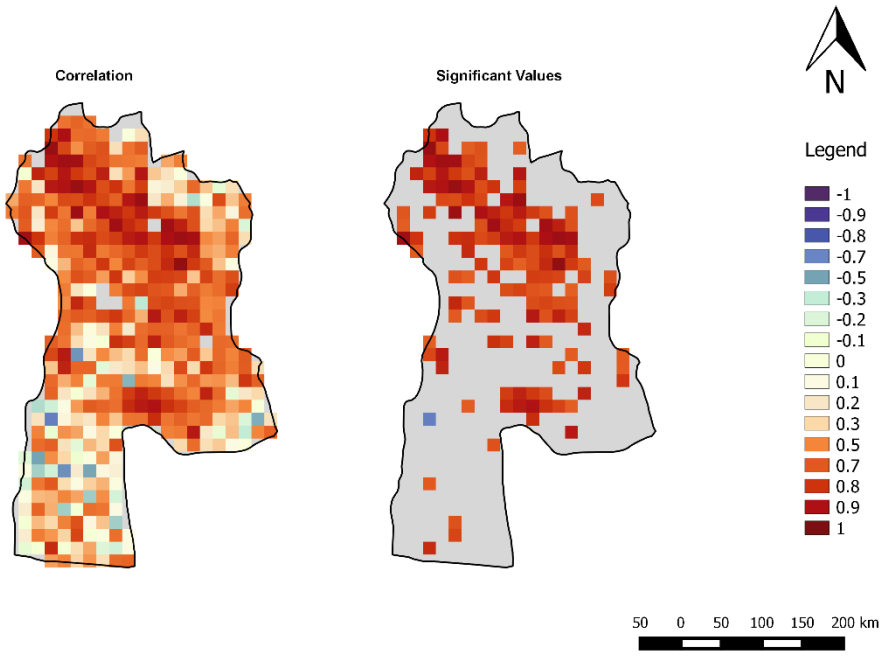


Figure 4.16 | Spatial distribution of temporal correlations between NDVI and Precipitation anomalies for the growing season of 2001. Significant values for p < 0,1.

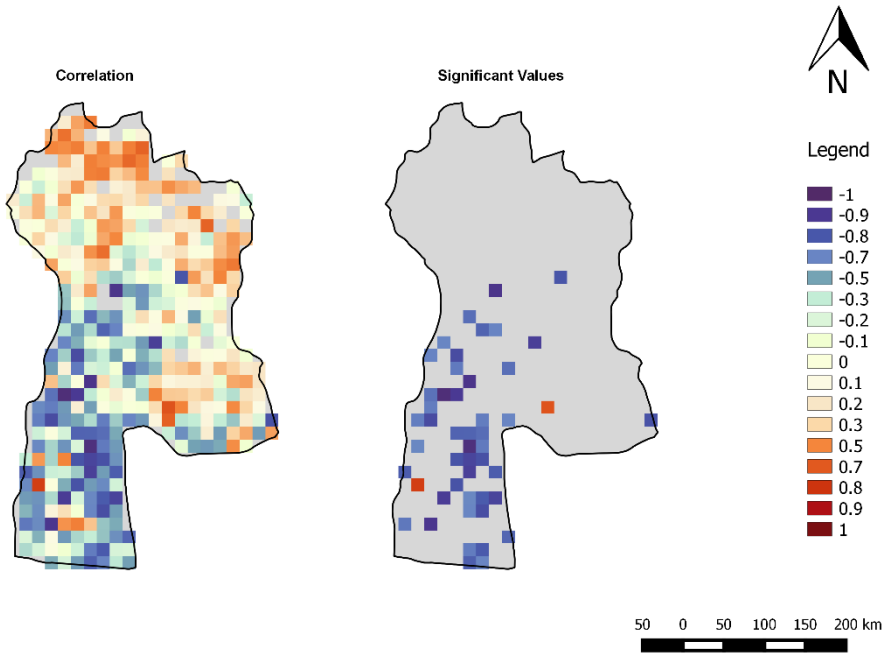


Figure 4.17 | Spatial distribution of temporal correlations between NDVI and Precipitation anomalies for the growing season of 2002. Significant values for p < 0,1.

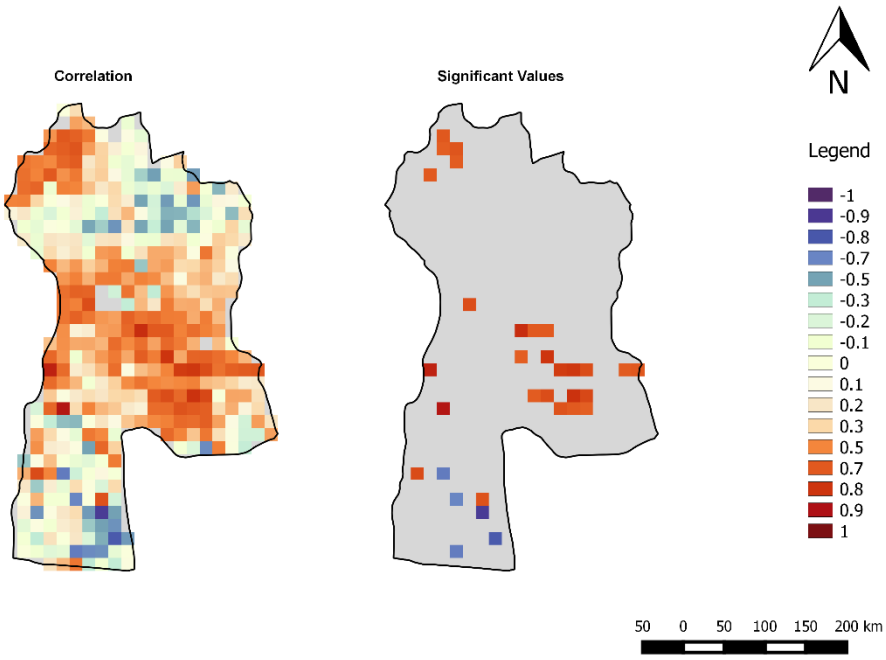


Figure 4.18 | Spatial distribution of temporal correlations between NDVI and Precipitation anomalies for the growing season of 2009. Significant values for p < 0,1.

5. Conclusions

The work presented in this study provided some temporal insight of the NDVI anomalies comparison with land surface temperature and precipitation anomalies in the region of Bayankhongor province, Mongolia. The study explored monthly MODIS NDVI, LST and TRMM Precipitation data for the growing season (March – September) over a period of fifteen years, focusing on known drought years (2001, 2002 and 2009). The aim of the study was to analyse the correlations between the variables anomalies and study the possibility of using NDVI as drought indicator.

Due to the high amount of files for each dataset and the pre-processing they had to go through, working primarily in QGIS software wasn't an option. Besides being a very time consuming process it would be more open to mistakes as a result of the repetitiveness. We used Python programming language to run algorithms to process all the images up to our goal. At the time I had no solid knowledge on programming so additional time was spent in learning and creating working programs for the study. However this led to a positive outcome which was the motivation to get a certificate (Appendix H) on Python programming language basics course from Michigan's University, through the well-known online platform Coursera[7].

The analysis of the monthly average for the three variables revealed that, generally, the NDVI reached higher values with high precipitation values and low LST values. While the year of 2001 was the driest year of the fifteen year period, the year of 2003 had the healthier vegetation activity. The last four years (2010-2014), show that NDVI trends have slightly increased as a result of higher precipitations and a small decrease of LST values.

The study of anomalies allowed a better understanding of the variables behaviour for each year of the time-series. In general, as expected, we may conclude that drought years exhibited the highest negative anomalies in NDVI and precipitation, and positive anomalies in LST. Monthly relationships between NDVI-LST and NDVI-precipitation anomalies for 2001, 2002 and 2009 revealed weak correlation results ($p<0,05$) with a few exceptions. The year of 2002 registered the strongest negative correlations between NDVI and LST anomalies for July ($r = -0,72$, $p < 0,05$) and August, ($r = -0,67$, $p < 0,05$), respectively. The positive correlations between NDVI and precipitation anomalies were also higher in July ($r = 0,43$, $p<0,05$) and August ($r = 0,52$, $p<0,05$) of 2002, as expected due to the summer rainfall patterns in the region. The year of 2001 revealed the weakest monthly correlations of the three years analysed.

Temporal correlation images for the growing season were generated to analyse the spatial distribution of the anomalies correlation. Since the monthly results consist in one single value, these images provided valuable information on which regions of Bayankhongor province had the strongest and weakest correlations. For NDVI-LST, the years of 2001 and 2002, showed negative

[7] Coursera.org has partnerships with the best universities and organizations around the world to provide universal online courses for free and each course has a paid option on which the student is evaluated and receives a certificate.

correlations in the north and positive correlations in the south. In 2001, the NDVI-Precipitation correlation registered strong values in the steppe and desert steppe region. However, in general the correlation values had little significance. Although precipitation in Mongolia occurs mostly during the growing season, the time lag between precipitation and NDVI should be taken into account for the next months. Further analysis should include this functional time lag in order to provide more robust conclusions.

The NDVI anomalies images analysis suggests that NDVI is a suitable drought index for Bayankhongor province, since the negative anomalies are well represented for the known drought years. However the weak overall correlations between the anomalies and the low significance of these values also suggests that an accurate answer for predicting drought using the relation between NDVI, LST and Precipitation cannot be given without additional research. Thus, correlation's studies should be deepen in areas with higher NDVI, since the sparse vegetation in the desert biomes may affect the regional results. The time-series used for this study focus on the last decade which include several extreme climatic events (droughts and *dzuds*) in the Mongolia region. Further studies should use a longer time-series in order to ensure a more reliable baseline and average values. Also, the use of a drought index for this region can ensure the data quality and validate NDVI values.

One important aspect to mention about this study is the focus on the methodological approach and data processing algorithms in order to achieve initial results about drought monitoring on the study site. Regardless of not being able to give more precise and conclusive results, this thesis provides a robust basis and a guide for additional studies and further research.

6. References

AghaKouchak, A., Farahmand, A., Melton, F. S., Teixeira, J., Anderson, M. C., Wardlow, B. D., & Hain, C. R. (2015). Remote sensing of drought: Progress, challenges and opportunities. *Reviews of Geophysics, 53*, 452-480.

Batima, P., Natsagdorj, L., Gombluudev, P., & Erdenetseteg, B. (2005). Observed climate change in Mongolia. *AIACC Working Paper, 12*, 12-16.

DEIMOS Engenharia. (2015). Internal project's report. *Earth Observation Support for Asian Development Bank Activities.* Lisbon.

Erch Partners. (2014, April 23). Bayankhongor aimag. Retrieved 7 10, 2015, from http://www.infomongolia.com/ct/ci/206/70/Bayankhongor%20Aimag4

Fondazione Edmund Mach (FEM). (2015). MODIS LST. *LST Details: QC maps*. S. Michele all'Adige (Trento), Italy. Retrieved June 18, 2015, from http://gis.cri.fmach.it/modis-lst/

Gopinath, G., Ambili, G., Gregory, S. J., & Anusha, C. (2014, June 29). Drought risk mapping of south-western state in the Indian peninsula - A web based application. *Journal of Environmental Management*, 161, 1-7.

Graham, S. (2000, August). Drought : The Creeping Disaster. United States. Retrieved August 20, 2015, from http://earthobservatory.nasa.gov/Features/DroughtFacts/

Ji, L., & Peters, A. J. (2003). Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. *Remote Sensing of Environment, 87*, 85-98.

John, R., Chen, J., Ou-Yang, Z.-T., Xiao, J., Becker, R., Samanta, A., Sandram, G., Wenping, Y. and Batkhishig, O. (2013). Vegetation response to extreme climate events on the Mongolian Plateau from 2000 to 2010. *Environmental Research Letters, 8*.

Karnieli, A., Agam, N., Pinker, R. T., Anderson, M., Imhoff, M. L., Gutman, G. G., Natalya ,P.and Goldberg, A. (2010). Use of NDVI and Land Surface Temperature for Drought Assessment: Merits and Limitations. *JOURNAL OF CLIMATE, 23*, 618-633.

Land Processes Distributed Active Archive Center (LP DAAC). (2014). Land Surface Temperature and Emissivity 8-Day L3 Global 1km. *MOD11A2*. Sioux Falls, South Dakota: DAAC, NASA EOSDIS Land Processes. Retrieved May 20, 2015, from https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod11a2

Land Processes Distributed Active Archive Center (LP DAAC). (2014, April 14). Vegetation Indices 16-Day L3 Global 250m. *MOD13Q1*. Sioux Falls, South Dakota, US: DAAC, NASA EOSDIS Land Processes. Retrieved May 20, 2015, from https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13q1

Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. (G. Syme, Ed.) *Journal of Hydrology*, 202-216.

Nandintsetseg, B., & Shinoda, M. (2012). Assessment of drought frequency, duration, and severity and its impact on pasture production in Mongolia. *Nat Hazards, 66*, 995-1008.

National Aeronautics and Space Administration, Goddard Space Flight Center. (2013). MODIS Grids. *Sinusoidal Tile Grid*. Retrieved June 26, 2015, from http://modis-land.gsfc.nasa.gov/MODLAND_grid.html

National Institute of Standards and Technology. (2008, March). Guide for the Use of the International System of Units (SI). *Degree Celsius, 2008*, 5 - 6. Gaithersburg: National Institute of Standards and Technology.

Nichol, J. E., & Abbas, S. (2015). Integration of remote sensing datasets for local scale assessment and prediction of drought. *Science of the Total Environment, 505*, 503-507.

Rhee, J., Im, J., & Carbone, G. J. (2010). Monitoring agricultural drought for arid and humid regions using multi-sensor. *Remote Sensing of Environment, 114*, 2875-2887.

Riebeek, H. (2009, September 4). Catalog of Earth Satellite Orbits. (E. Observatory, Ed.) NASA, USA. Retrieved September 10, 2015, from http://earthobservatory.nasa.gov/Features/OrbitsCatalog/

Senay, G. B., Velpuri, N. M., Bohms, S., Budde, M., Young, C., Rowland, J., & Verdin, J. P. (2015). Drought Monitoring and Assessment: Remote Sensing and Modeling Approaches for the Famine Early Warning Systems Network. In *Hydro-Meteorological Hazars, Risks and Disasters* (pp. 233–262). Elsevier.

Son, N. T., Chen, C. F., Chen, C. R., Chang, L. Y., & Minh, V. Q. (2012, March). Monitoring agricultural drought in the Lower Mekong Basin using MODIS NDVI and land surface temperature data. *International Journal of Applied Earth Observation and Geoinformation*, 417-427.

Sruthi, S., & Mohammed Aslam, M. A. (2015). Agricultural Drought Analysis Using the NDVI and Land Surface Temperature Data; a Case Study of Raichur District. *INTERNATIONAL CONFERENCE ON WATER RESOURCES, COASTAL AND OCEAN ENGINEERING (ICWRCOE 2015)* (pp. 1258 – 1264). Elsevier.

Sternberg, T., Thomas, D., & Middleton, N. (2011). Short Communication Drought dynamics on the Mongolian steppe, 1970–2006. *International Journal of Climatology, 31*, 1823-1830.

The World Bank Group. (2015). Mongolia Dashboard. *Climate Baseline*. Retrieved 8 20, 2015, from

http://sdwebx.worldbank.org/climateportalb/home.cfm?page=country_profile&CCode=MNG&ThisTab=ClimateBaseline

Tropical Rainfall Measuring Mission Project (TRMM). (2011). TRMM and Other Sources Monthly Rainfall Product (TRMM Product 3B43 v7). *version 7*. Greenbelt, MD:Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC),. Retrieved June 12, 2015, from http://disc.sci.gsfc.nasa.gov/datacollection/TRMM_3B43_V7.html

Tucker, C. J. (1979, May 1). Red and photographic infrared linear combinations for monitoring vegetation. (U. States, Ed.) *Remote Sensing of Environment, 8*(EARTH RESOURCES AND REMOTE SENSING), 127-150. doi:10.1016/0034-4257(79)90013-0

UNEP. (2002). Overview of major environmental developments and trends. In *Mongolia : State of the Environment 2002.* (pp. 9-50). United Nations Environment Programme.

Wardlow, B. D., Anderson, M. C., & Verdin, J. P. (2012). *Remote Sensing of Drought: Innovative Monitoring Approaches.* United States of America: Taylor & Francis Group, LLC.

Weier, J., & Herring, D. (2000). Measuring Vegetation (NDVI & EVI). NASA Earth Observatory. Retrieved August 20, 2015, from http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_1.php

Weier, J., & Herring, D. (2000, 8 30). Measuring Vegetation (NDVI and EVI). *Normalized Difference Vegetation Index*. (N. E. (EO), Ed.) Retrieved 3 20, 2015, from http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_1.php

Zhang, L., Xiao, J., Li, J., Wang, K., Lei, L., & Guo, H. (2012). The 2010 spring drought reduced primary productivity in southwestern China. *Environmental Research Letters, 7*.

ХЭЛТЭС, БАЯНХОНГОР АЙМАГ СТАТИСТИКИЙН. (2014). 2014 оны хүн амын үндсэн үзүүлэлтүүд. *Тооллого судалгаа*. (Y. C. ХОРОО, Ed.) Bayankhongor, Mongolia. Retrieved August 2, 2015, from http://bayankhongor.nso.mn/uploads/users/7/files/HUN%20AMIIN%20UNDSEN%20UZUULELT_BHR.pdf

7. Appendixes

Appendix A – Python programs

Appendix A.1 Convert HDF to TIF (LST)

```
# -*- coding: utf-8 -*-
"""
@author: JFG
"""
#Extract selected SDS from HDF File (convert and reproject) and mosaic tiles
based on namefile patterns, pymodis and gdal lib (batch)

import os,fnmatch
import os.path
import sys
sys.path.append('C:/OSGeo4W64/bin')
from pymodis import convertmodis_gdal
import gdal_merge
from subprocess import call

#folders
Appendix     inputFolder="C:/Users/Joana/Documents/Tese_Computador/Datasets_Or
   iginais_TESE/Temperatura_O/MOD11A2/Todos"

#pattern of the hdf files
pattern_hdf='MOD11A2'+'*'+'.[Hh][Dd][Ff]'

#pattern for LST tif files
pattern_24='MOD11A2'+'*'+'h24'+'*'+'LST_Day_1km'+'.[Tt][Ii][Ff]'
pattern_25='MOD11A2'+'*'+'h25'+'*'+'LST_Day_1km'+'.[Tt][Ii][Ff]'

#Lists with file names
LN24 = []
LN25 = []
#Lists with paths for the files
LP24 = []
LP25 = []

#-----------------------------------------------------------------#
#------Extracts, Converts and Reprojects layers from HDF-------#
#-----------------------------------------------------------------#

#parameters for convertmodis
#hdfname =                          #(str) - name of input data
#prefix =                           #(str) - prefix for output data
subset = '1 1'                      #(str) - the subsets (LSD_Day and QC_day)
res = 1000                          #(int) - output resolution
outformat = 'Gtiff'                 #(str) - output format
epsg = 32647                        #(int) - output EPSG code
wkt = None                          #(str) - output WKT string
resampl = 'NEAREST_NEIGHBOR'        #(str) - the resampling method
vrt = False                         #(bool)- True to read GDAL VRT file
created with createMosaicGDAL


#enters the input folder and "walks" through all the files in it
for root, dirs, files in os.walk(inputFolder):
    #browse all files that match the pattern
    for hdfname in fnmatch.filter(files,pattern_hdf):
        a = hdfname.find('.hdf')
        #slice the name of each hdf file to use it as prefix
        prefix = hdfname[:a]
        #Extracts and converts layers to Gtiff
        convert = convertmodis_gdal.convertModisGDAL(hdfname, prefix, subset,
res, outformat, epsg, wkt, resampl, vrt)
```

```
        convert.run()


#-----------------------------------------------------------------#
#--------------------Merge Tif files------------------------------#
#-----------------------------------------------------------------#

#enters the input folder and "walks" through all the files in it
for root, dirs, files in os.walk(inputFolder):

#browse all files that match the patterns

    #searches for all the "h" LST tif files and ads their names to the LN
lists and their paths to the LP list
    for tif24 in fnmatch.filter(files,pattern_24):

            p1 = os.path.realpath(os.path.join(root,tif24))
            LP24.append(p1)
            LN24.append(tif24)

    for tif25 in fnmatch.filter(files,pattern_25):

             p2 = os.path.realpath(os.path.join(root,tif25))
             LP25.append(p2)
             LN25.append(tif25)


#Applies the merge function to each pair of LST files (h24 + h25)
#When it reaches the Lists lenght it stops the "for" loop because there's no
more elements to add to the function
for i in range(len(LP24)):
    print "i e j"
    print LP24[i]
    print LP25[i]

    #Converts each element (string) in the list to a path before going in the
gdal_merge function
    inputTif1 = os.path.abspath(LP24[i])
    inputTif2 = os.path.abspath(LP25[i])

    print inputTif1
    print inputTif2

    #gdal_merge arguments and main function
    sys.argv = ['-o','outputTif12','-of','GTiff','-ot','Float64','-n','0','-
a_nodata', '-9999',inputTif1,inputTif2]
    gdal_merge.main()

    #output name
    name_output
=str('MOD11A2'+LN24[i][8:16]+'_h24h25v04_005_'+LN24[i][28:53]+'.tif')

    print name_output

    #Renames the out.tif to the name_output
    call(['gdalmanage', 'rename', 'out.tif', name_output])
    call(['gdalmanage', 'delete', 'out.tif'])

print "Done"
```

Appendix A.2 Convert NC to TIF, Clip and Calculate Mean (Precipitation)

```
# -*- coding: utf-8 -*-
"""
@author: JFG
"""
#Extract selected SDS from HDF File (convert and reproject) and mosaic
tiles based on name file patterns, pymodis and gdal lib (batch)

from osgeo import gdal
from osgeo.gdalconst import *
import os,fnmatch
import os.path
import sys
sys.path.append('C:/OSGeo4W64/bin')
from subprocess import call
import numpy as np

#folders
inputFolder="C:/Users/Joana/Documents/Tese_Computador/Datasets_Origina
is_TESE/Precipitacao_O/TRMM_3B43/MarOut_P/1_2_Marco"

#pattern of the hdf files
pattern_nc='3B43'+'*'+'.[Nn][Cc]'
#pattern_hdf='3B43'+'*'+'.[Hh][Dd][Ff]'
pattern_tif2='3B43'+'*'+'_1.[Tt][Ii][Ff]'

#-----------------------------------------------------------------#
#-----------Converts and Reprojects NC to TIFF----------------#
#-----------------------------------------------------------------#

#browses all files that match the pattern inside the input folder and
runs convert modis for each one (extracts and converts layers to
Gtiff)
for root, dirs, files in os.walk(inputFolder):

    for ncname in fnmatch.filter(files,pattern_nc):
        a = ncname.find('.nc')
        prefix = ncname[:a-3] #slice the name of each nc file to use
it as prefix without ending in .nc
        print ncname

        print ' EXTRACTING'
        #gdal_translate converts raster data between different formats
        call(["gdal_translate", "-ot", "Float64","-of", "GTiff","-
co","TILED=YES","-a_srs","EPSG:4326","-sds","-a_nodata","-9999",
root+"/"+ncname, root+"/"+prefix+".tif"])

for root, dirs, files in os.walk(inputFolder):
    for tifname in fnmatch.filter(files,pattern_tif2):
        b = tifname.find('.tif')
        prefix2 = tifname[:b] #slice the name of each hdf file to use
it as prefix without ending in .tif
        print "TIFF NAME"
        print tifname
        #Path for the shape file used to clip the original image
        clip =
"C:/Users/Joana/Documents/Tese_Computador/Datasets_Originais_TESE/Baya
nkhongor/Bayankhongor_reproj.shp"

        print '----> Reprojecting'
```

```
        #gdalwarp allows the reprojection of the tif file generated by
gdal_translate (to EPSG:32647)
        call(["gdalwarp", "-t_srs", "EPSG:32647","-of", "GTiff","-
co","TILED=YES",root+"/"+tifname,root+"/"+"R_"+prefix2+".tif"])

        print '----> CLIPPING'
        call(["gdalwarp", "-cutline", clip,"-crop_to_cutline","-
overwrite", "-of", "GTiff","-
co","TILED=YES",root+"/"+"R_"+prefix2+".tif",
root+"/"+"CLIPR_"+prefix2+".tif"])

print '----> AVERAGE'
#---------------AVERAGE-------------------------#
#pattern of the tif files (_1: precipitation layer)
pattern_tif='CLIPR'+'*'+'.[Tt][Ii][Ff]'
pattern_avgtif='AVERAGE_*.[Tt][Ii][Ff]'
pattern_counttif='COUNT_*.[Tt][Ii][Ff]'

#-------------------------------------------------#
#----------- Function of development ------------#
#-------------------------------------------------#

def calc_average_Landsat(allarrays):


   meanarray = np.nanmean(allarrays,axis=0)
   return meanarray


#-------------------------------------------------#
#--------------- Apply Function -----------------#
#-------------------------------------------------#

#Creates a list with all the tif files that match the pattern
for root, dirs, files in os.walk(inputFolder):

    for tifname in fnmatch.filter(files,pattern_tif):
        print tifname
        tif = gdal.Open(root+"/"+tifname,gdal.GA_ReadOnly)
        arraytmp =
[[np.array(tif.GetRasterBand(1).ReadAsArray(),dtype=float)]]
        x_exists = 'allarrays' in locals() or 'allarrays' in globals()
        if x_exists:
            teste=np.reshape(arraytmp[0][0],-1)
            allarrays=np.append(allarrays,[teste],axis=0)
            print np.shape(allarrays)

        else:
            print np.shape(arraytmp[0][0])
            teste=np.reshape(arraytmp[0][0],-1)
            print np.shape(teste)
            allarrays=[teste]

            print np.shape(allarrays)
        tif = None


print np.shape(allarrays)
boolarray=allarrays!=0
validpointsarray=np.sum(boolarray,axis=0)
allarrays[allarrays==-9999]=np.nan
```

```
meanarray = calc_average_Landsat(allarrays)

print np.shape(meanarray)

tifname1=fnmatch.filter(files,pattern_avgtif)
tifname2=fnmatch.filter(files,pattern_counttif)
print tifname
if tifname:
    avgtif = gdal.Open(root+"/"+tifname1[0],gdal.GA_Update)
    counttif = gdal.Open(root+"/"+tifname2[0],gdal.GA_Update)
    cols = avgtif.RasterXSize
    rows = avgtif.RasterYSize
    print rows, cols
    meanarray=np.reshape(meanarray,(rows,cols))
    validpointsarray=np.reshape(validpointsarray,(rows,cols))
    print np.shape(meanarray)
    avgtif.GetRasterBand(1).WriteArray(meanarray)
    counttif.GetRasterBand(1).WriteArray(validpointsarray)
    avgtif = None
    counttif = None

print "Done!"
```

Appendix A.3 Clip Bayankhongor province from raster

```
# -*- coding: utf-8 -*-
"""
@author: JFG
"""
#Clip Raster
#See: <http://gis.stackexchange.com/questions/75086/how-to-for-loop-a-folder-to-batch-clip-rasters-by-polygon-using-python-and-qgis

import os, fnmatch
from subprocess import call

#Folders
inFolder=
'C:/Users/Joana/Documents/Tese_Computador/Datasets_Originais_TESE/Temperatura_O/MOD11A2/Mar_Out_T/TIFS_T/'
outFolder=
'C:/Users/Joana/Documents/Tese_Computador/Datasets_Originais_TESE/Temperatura_O/MOD11A2/Mar_Out_T/TIFS_T/'

#Pattern of tif files
pattern_tif= 'MOD11A2'+'*'+'h24h25v04'+'*'+'.[Tt][Ii][Ff]'
os.chdir (inFolder)

#enters the input folder and "walks" through all the files in it
for root, dirs, files in os.walk(inFolder):

    #browse all files that match the pattern
    for tifname in fnmatch.filter(files,pattern_tif):
        b = tifname.find('.tif')
        #slice the name of each hdf file to use it as prefix
        prefix2 = tifname[:b]
        print "TIFF NAME"
        print tifname
        #Path for the shape file we want to use to clip the original image (Bayahnkhongor shape file)
        clip =
"C:/Users/Joana/Documents/Tese_Computador/Datasets_Originais_TESE/Bayankhongor/Bayankhongor_reproj.shp"
        #Apply gdalwarp
        call(["gdalwarp", "-cutline", clip,"-crop_to_cutline","-overwrite", "-of", "GTiff","-co","TILED=YES",root+tifname, root+"/"+"C_"+prefix2+".tif"])

print 'Done CLIP!'
```

Appendix A.4 Multiply by scale factor and apply QC mask

```
# -*- coding: utf-8 -*-
"""
@author: capp
Modified: Joana
"""

from osgeo import gdal
from osgeo.gdalconst import *
import os,fnmatch
import numpy as np
from subprocess import call


#-----------------------------------------------------#
#---Choose the input folder and pattern tiff/qcmask---#
#-----------------------------------------------------#

#folders
inputFolder="C:/Users/Joana/Documents/Tese_Computador/Datasets_Originais_TESE/
Temperatura_O/MOD11A2/Mar_Out_T/CLIP_T/"

# Scale Factor to multiply by original values
scale_factor = 0.02

#-------------------------------------------------#
#----------- Function of development ------------#
#-------------------------------------------------#

def apply_mask_Landsat(root, filename_mask, filename_tiff, scale_factor):


    #Open the qcmask and create a array file with the dtype=float
    qcmask = gdal.Open(root+"/"+filename_mask,gdal.GA_Update)

    array_qcmask = np.array(qcmask.GetRasterBand(1).ReadAsArray(),dtype=float)

    #Name Out for the new file tiff with apply mask

    name_out ="MASK_"+filename_tiff

    # Call the gdal_translate because it needs to be same type of the float
    call(["gdal_translate", "-ot", "Float64","-of", "GTiff","-co",
"TILED=YES", root+"/"+filename_tiff, root+"/"+name_out])

    qcmask =None

    # Open same created file and convert to a array also type float
    new_band=gdal.Open(root+"/"+name_out,gdal.GA_Update)

    array_band=np.array(new_band.GetRasterBand(1).ReadAsArray(),dtype=float)

    nodatavalue = new_band.GetRasterBand(1).GetNoDataValue()

    print nodatavalue
    array_band[array_band != nodatavalue]

    # Multiplication of the original values by the scale factor and conversion
from kelvins to celsius
    array_band[array_band != nodatavalue] =
(scale_factor*array_band[array_band != nodatavalue])-273.15

    # Change all values where qcmask has bad quality pixels and transform them
in values without data = -9999
    # http://gis.cri.fmach.it/modis-lst/ See table at LST Details: QC maps
    array_band[np.logical_or(array_qcmask == 2,array_qcmask == 3,array_qcmask
>= 129)] = -9999
```

```
    # Write array write data on the new tiff in float 64
    new_band.GetRasterBand(1).WriteArray(array_band)

    new_band = None


#-----------------------------------------------#
#---------------- Apply Function----------------#
#-----------------------------------------------#

#pattern of the tif files
pattern_tiff='C_'+'*'+'h24h25'+'*'+'LST_Day_1km'+'.[Tt][Ii][Ff]'
pattern_mask='C_'+'*'+'h24h25'+'*'+'QC_Day'+'.[Tt][Ii][Ff]'

L1 = []
L2 = []

print 'first FOR'
#enter the input folder and "walks" through all the files in it
for root, dirs, files in os.walk(inputFolder):

#compares the files inside the input folder to the patterns
    for filename_mask1 in fnmatch.filter(files, pattern_mask):
        print filename_mask1
        L1.append(filename_mask1)

    for filename_tiff1 in fnmatch.filter(files, pattern_tiff):
        print filename_tiff1
        L2.append(filename_tiff1)


print 'Second FOR'

for i in range(len(L1)):
    print "i e j"
    print L1[i]
    print L2[i]

    #Converts each element (string) in the list to a path before going in the
gdal_merge function
    filename_mask = L1[i]
    filename_tiff = L2[i]

    print 'FILENAMES'
    print filename_mask
    print filename_tiff

    apply_mask_Landsat(root, filename_mask, filename_tiff, scale_factor)


print "Done"
```

Appendix A.5 Monthly average

```
# -*- coding: utf-8 -*-
"""
@author: JFG
Help: NG
"""
from osgeo import gdal
from osgeo.gdalconst import *
import os,fnmatch
import numpy as np
from subprocess import call
from itertools import izip

#calculates average between tiff files
#-------------------------------------------------------#
#-------Choose the input folder and pattern tiff------#
#-------------------------------------------------------#

#folders
inputFolder="C:/Users/Joana/Documents/Tese_Computador/Datasets_Originais_TESE/
Temperatura_O/MOD11A2/Mar_Out_T/LST_Day/1_Marco"

#year input
year = raw_input("What year?")

#pattern of the tif files
pattern_tif='MASK_C_MOD11A2A'+year+'*'+'.[Tt][Ii][Ff]'
pattern_avgtif='AVERAGE_'+'*'+year+'*'+'.[Tt][Ii][Ff]'
pattern_counttif='COUNT_'+'*'+year+'*'+'.[Tt][Ii][Ff]'

#-------------------------------------------------#
#----------- Function of development ------------#
#-------------------------------------------------#

def calc_average_Landsat(allarrays):

   #function to calculate the mean between the axis without considering the
nodatavalue
   meanarray = np.nanmean(allarrays,axis=0)
   return meanarray


#-------------------------------------------------#
#--------------- Apply Function -----------------#
#-------------------------------------------------#

#Creates a list with all the tiff files that match the pattern
for root, dirs, files in os.walk(inputFolder):

    for tifname in fnmatch.filter(files,pattern_tif):
        tif = gdal.Open(root+"/"+tifname,gdal.GA_ReadOnly)
        arraytmp =
[[np.array(tif.GetRasterBand(1).ReadAsArray(),dtype=float)]]
        print tifname
        x_exists = 'allarrays' in locals() or 'allarrays' in globals()
        if x_exists:
            teste=np.reshape(arraytmp[0][0],-1)
            allarrays=np.append(allarrays,[teste],axis=0)
            print np.shape(allarrays)

        else:
            print np.shape(arraytmp[0][0])
            teste=np.reshape(arraytmp[0][0],-1)
            print np.shape(teste)
            allarrays=[teste]

            print 'SHAPE ALL ARRAYS'
```

```
            print np.shape(allarrays)
        tif = None


print np.shape(allarrays)
boolarray=allarrays!=0
validpointsarray=np.sum(boolarray,axis=0)
allarrays[allarrays==-9999]=np.nan

meanarray = calc_average_Landsat(allarrays)

print 'SHAPE MEANARRAY'
print np.shape(meanarray)

tifname1=fnmatch.filter(files,pattern_avgtif)
tifname2=fnmatch.filter(files,pattern_counttif)
print tifname1
if tifname:
    avgtif = gdal.Open(root+"/"+tifname1[0],gdal.GA_Update)
    counttif = gdal.Open(root+"/"+tifname2[0],gdal.GA_Update)
    cols = avgtif.RasterXSize
    rows = avgtif.RasterYSize
    #print rows, cols
    meanarray=np.reshape(meanarray,(rows,cols))
    validpointsarray=np.reshape(validpointsarray,(rows,cols))
    print np.shape(meanarray)
    #Updates the values on the existing files to match the meanarray
    avgtif.GetRasterBand(1).WriteArray(meanarray)
    counttif.GetRasterBand(1).WriteArray(validpointsarray)
    avgtif = None
    counttif = None
    print "Done!"
```

Appendix A.6 Anomaly calculation

```
# -*- coding: utf-8 -*-
"""
@author: JFG
"""
from osgeo import gdal
from osgeo.gdalconst import *
import os,fnmatch
import numpy as np
from subprocess import call
from itertools import izip


#-------------------------------------------------------#
#---Choose the input folder and pattern tiff----------#
#-------------------------------------------------------#

#folders
inputFolder="C:/Users/Joana/Documents/Tese_Computador/Datasets_Origina
is_TESE/NDVI_O/MarOut_NDVI/2_Abril"

#pattern of the tif files
pattern_tif='AVERAGE_APRIL'+'*'+'.[Tt][Ii][Ff]'
pattern_tif2='ANOMALY'+'*'+'.[Tt][Ii][Ff]'

#-------------------------------------------------#
#----------- Function of development ------------#
#-------------------------------------------------#

def calc_anomaly(filename_1):
    print 'FILENAME INSIDE FUNCTION'
    print filename_1

    #root to mean file
    allmeans =
"C:/Users/Joana/Documents/Tese_Computador/Datasets_Originais_TESE/NDVI
_O/MarOut_NDVI/2_Abril/AVERAGE_ALL_April_2000_2014_NDVI_BAYA_MOD13Q1.t
if"

    # Open the input files and create a array file with the
dtype=float
    input1 = gdal.Open(root+"/"+filename_1,gdal.GA_Update)
    input2 = gdal.Open(allmeans,gdal.GA_Update)

    array_input1 =
np.array(input1.GetRasterBand(1).ReadAsArray(),dtype=float)
    array_input2 =
np.array(input2.GetRasterBand(1).ReadAsArray(),dtype=float)
    np.append(array_input1,array_input2,axis=0)

    # Output name for tif file with calculated average
    name_out ="ANOMALY_"+filename_1
    print 'Output FILE:'
    print name_out

    # Call the gdal_translate because it needs to be same type of the
float
    call(["gdal_translate", "-ot", "Float64","-of", "GTiff","-co",
"TILED=YES", root+"/"+filename_1, root+"/"+name_out])

    input1 = None
```

```
    input2 = None

    # Open same created file and convert to a array also type float
    new_band= gdal.Open(root+"/"+name_out,gdal.GA_Update)
    array_band=
np.array(new_band.GetRasterBand(1).ReadAsArray(),dtype=float)
    nodatavalue = new_band.GetRasterBand(1).GetNoDataValue()

    #print nodatavalue
    #print np.shape(array_band)
    #print np.shape(array_input1)
    #print np.shape(array_input2)
    array_input_all=[array_input1,array_input2]

    array_band[(array_input1 != nodatavalue) & (array_input2 !=
nodatavalue)] = (array_input1[(array_input1 != nodatavalue) &
(array_input2 != nodatavalue)] - array_input2[(array_input1 !=
nodatavalue) & (array_input2 != nodatavalue)])

    # Write array write data on the new tiff in float 64
    new_band.GetRasterBand(1).WriteArray(array_band)

    new_band = None

#-------------------------------------------------#
#--------------- Apply Function ------------------#
#-------------------------------------------------#
#Creates a list with all the tif files that match the pattern
for root, dirs, files in os.walk(inputFolder):

    for tifname in fnmatch.filter(files,pattern_tif):
        print 'TIF NAME'
        print tifname
        calc_anomaly(tifname)
#-------------------------------------------------#
#------------ SAVE DATA TO CSV -------------------#
#-------------------------------------------------#
L1 =  list()
for root, dirs, files in os.walk(inputFolder):

    for tifname in fnmatch.filter(files,pattern_tif2):

        a = tifname.find(".tif")
        prefix = tifname[:a-3]

        filename = gdal.Open(root+"/"+tifname,gdal.GA_Update)
        print tifname
        tif1 =
[[np.array(filename.GetRasterBand(1).ReadAsArray(),dtype=float)]]

        #Gets the no data value for each file
        nodatavaluetif = filename.GetRasterBand(1).GetNoDataValue()

        filename = None
        n = np.asarray(tif1)

        print 'SHAPES'
        print np.shape(n)

        n = np.reshape(n[0][0],-1)
        print 'no data:'
```

```
            print nodatavaluetif

            n = n[np.logical_not(np.isnan(n))]
            #print n[0]
            #print type(n[0])

            n = n[n != nodatavaluetif]
            print np.shape(n)
            m = np.average(n)
            L1.append(m)

np.savetxt(prefix+".csv", L1, delimiter=",")

print "Done!"
```

Appendix A.7 Resample to coarser resolution

```
# -*- coding: utf-8 -*-
"""
@author: Joana
"""
#Reproject raster images
from osgeo import gdal
from osgeo.gdalconst import *
import os,fnmatch
import os.path
import sys
sys.path.append('C:/OSGeo4W64/bin')
from subprocess import call
import numpy as np

#folders
inputFolder="C:/Users/Joana/Documents/Tese_Computador/Datasets_Originais_TESE/Temperatura_O/MOD11A2/Mar_Out_T/LST_Day/1_Marco"

#pattern of the tif files
pattern_tif='ANOMALY'+'*'+'.[Tt][Ii][Ff]'

#-----------------------------------------------#
#----------------Resample tifs------------------#
#-----------------------------------------------#

#browse all files that match the pattern inside the input folder

for root, dirs, files in os.walk(inputFolder):
    for tifname in fnmatch.filter(files,pattern_tif):
        b = tifname.find('.tif')
        prefix2 = tifname[:b] #slice the name of each hdf file to use it as prefix without ending in .tif
        print "TIFF NAME"
        print tifname

        #gdalwarp allows the resampling of the tif file to pixel size of 15935.5
        call(["gdalwarp", "-t_srs", "EPSG:32647","-tr","15935.5","15894.3","-of", "GTiff","-co","TILED=YES",root+"/"+prefix2+".tif",root+"/"+"R_"+prefix2+".tif"])
        #confirm new image shape to see if matches Precipitation images shape (36,21)
        c = root+"/"+"R_"+prefix2+".tif"
        tif = gdal.Open(c,gdal.GA_ReadOnly)
        arraytmp = [[np.array(tif.GetRasterBand(1).ReadAsArray(),dtype=float)]]
        print 'SHAPE'
        print np.shape(arraytmp[0][0])

print "Done"
```

Appendix A.8 Pearson Correlation

```
# -*- coding: utf-8 -*-
"""
@author: JFG
Help: NG
"""
import sys
import csv
from osgeo import gdal
from osgeo.gdalconst import *
import os,fnmatch
import numpy as np
from subprocess import call
from scipy.stats.stats import pearsonr

#Folder where the files are
inputFolder=
'C:/Users/Joana/Documents/Tese_Computador/Datasets_Originais_TESE/Corr
elacoes'

month= raw_input('Month: ')
year= raw_input('Year: ')
#patterns for the tifs we want to correlate
pattern_ndvi='R_'+'*'+month+'*'+'MOD13Q1'+'*'+year+'*'+'.[Tt][Ii][Ff]'
#pattern_temp='R_'+'*'+'MOD11A2'+'*'+month+year+'*'+'.[Tt][Ii][Ff]'
pattern_prec='ANOMALY_'+month+'*'+'CLIPR_3B43.'+year+'*'+'.[Tt][Ii][Ff
]'

#-----------------------------------------------------#
#-------------Apply CORRCOEF function-----------------#
#-----------------------------------------------------#

#Goes through all the files in the inputfolder
for root, dirs, files in os.walk(inputFolder):

    #identifies the tifs that match the patterns
    tifname1=fnmatch.filter(files,pattern_ndvi)
    #tifname2=fnmatch.filter(files,pattern_temp)
    tifname3=fnmatch.filter(files,pattern_prec)
    print tifname1
    print tifname3
    ndvitif = gdal.Open(root+"/"+tifname1[0],gdal.GA_Update)
    print 'ndvi tif array'
    #temptif = gdal.Open(root+"/"+tifname2[0],gdal.GA_Update)
    prectif = gdal.Open(root+"/"+tifname3[0],gdal.GA_Update)
    #info about the rows/cols in the tif files
    print 'NDVI'
    cols = ndvitif.RasterXSize
    rows = ndvitif.RasterYSize
    print rows, cols
    """print 'TEMPERATURE'
    cols2 = temptif.RasterXSize
    rows2 = temptif.RasterYSize
    print rows2, cols2
    """
    print 'PRECIPITATION'
    cols3 = prectif.RasterXSize
    rows3 = prectif.RasterYSize
    print rows3, cols3

print ' Creating array with tif files info'
```

```
#Creates arrays witht the tif files info
ndvi =
[[np.array(ndvitif.GetRasterBand(1).ReadAsArray(),dtype=float)]]
#temp =
[[np.array(temptif.GetRasterBand(1).ReadAsArray(),dtype=float)]]
prec =
[[np.array(prectif.GetRasterBand(1).ReadAsArray(),dtype=float)]]

print ' Getting nodata value'
#Gets the no data value for each file
nodatavaluendvi = ndvitif.GetRasterBand(1).GetNoDataValue()
#nodatavaluetemp = temptif.GetRasterBand(1).GetNoDataValue()
nodatavalueprec = prectif.GetRasterBand(1).GetNoDataValue()

ndvitif = None
#temptif = None
prectif = None

n = np.asarray(ndvi)
#t = np.asarray(temp)
p = np.asarray(prec)

print 'SHAPES (NDVI, TEMP, PREC)'
print np.shape(n)
#print np.shape(t)
print np.shape(p)

n = np.reshape(n[0][0],-1)
#t = np.reshape(t[0][0],-1)
p = np.reshape(p[0][0],-1)

print np.shape(n)
print np.shape(p)
print nodatavaluendvi, nodatavalueprec

#the indexes allow to discart the nodata values (which would bring the
corr result up)
#idx1 = np.logical_and(np.logical_not(np.isnan(n)) ,
np.logical_not(np.isnan(t)))
idx2 = np.logical_and(n != nodatavaluendvi , p != nodatavalueprec)

#n = n[idx1]
#t = t[idx1]
n = n[idx2]
p = p[idx2]
np.savetxt(tifname1[0]+".csv", n, delimiter=",")
#np.savetxt(tifname2[0]+".csv", t, delimiter=",")
np.savetxt(tifname3[0]+".csv", p, delimiter=",")

#print 'SHAPES 2 (NDVI, TEMP, PREC)'
#print np.shape(n)
#print np.shape(t)
#print np.shape(p)

print 'CORRCOEF RESULT:'
resultcor = np.corrcoef(n,p)
print resultcor
print 'PEARSONR RESULT:'
resultpearson = pearsonr(n,p)
print resultpearson
print 'Done!'
```

Appendix A.9 Pixelwise correlation between datasets

```
# -*- coding: utf-8 -*-
"""
@author: JFG
"""
from osgeo import gdal
from osgeo.gdalconst import *
import os,fnmatch
import numpy as np
from subprocess import call
from scipy import stats
from scipy.stats.stats import pearsonr

# Goal: Get correlation values for each pixel between two long term
datasets

#-----------------------------------------------------#
#---Choose the input folder and pattern tiff----------#
#-----------------------------------------------------#

#folders
inputFolder="C:/Users/Joana/Desktop/Testes_py/correlacao"

#pattern of the tif files
pattern_ndvi='R_ANOMALY'+'*'+'NDVI'+'*'+'.[Tt][Ii][Ff]'
pattern_temp='R_ANOMALY'+'*'+'MOD11A2'+'*'+'.[Tt][Ii][Ff]'
#pattern_prec='R_ANOMALY'+'*'+'.[Tt][Ii][Ff]'

# Create arrays with
print 'Creating NDVI array'
for root, dirs, files in os.walk(inputFolder):

    #identifies the tifs that match the pattern
    for tifname1 in fnmatch.filter(files,pattern_ndvi):
        #print tifname1
        ndvitif = gdal.Open(root+"/"+tifname1,gdal.GA_Update)
        ndvi =
[[np.array(ndvitif.GetRasterBand(1).ReadAsArray(),dtype=float)]]
        #print np.shape(ndvi)
        #print ndvi[1]

        x_exists = 'allarraysndvi' in locals() or 'allarraysndvi' in
globals()
        if x_exists:
            teste=np.reshape(ndvi[0][0],-1)
            allarraysndvi=np.append(allarraysndvi,[teste],axis=0)
            np.savetxt(tifname1+".csv", teste, delimiter=",")
            #print np.shape(allarrays)

        else:
            #print np.shape(ndvi[0][0])
            teste=np.reshape(ndvi[0][0],-1)
            #print np.shape(teste)
            allarraysndvi=[teste]
        #print np.shape(allarraysndvi)
        ndvitif= None

allarraysndvi[allarraysndvi==-9999]=np.nan
np.savetxt("allarrays"+".csv", allarraysndvi, delimiter=",")
#print allarraysndvi[0,5]
#print len(allarraysndvi)
```

```
#print len(allarraysndvi[0])

print 'Creating Temperature array'
for root, dirs, files in os.walk(inputFolder):

    for tifname2 in fnmatch.filter(files,pattern_temp):
        #print tifname2
        temptif = gdal.Open(root+"/"+tifname2,gdal.GA_Update)
        temp =
[[np.array(temptif.GetRasterBand(1).ReadAsArray(),dtype=float)]]
        x_exists = 'allarraystemp' in locals() or 'allarraystemp' in
globals()
        if x_exists:
            teste2=np.reshape(temp[0][0],-1)
            allarraystemp=np.append(allarraystemp,[teste2],axis=0)
            #print np.shape(allarrays)

        else:
            #print np.shape(ndvi[0][0])
            teste2=np.reshape(temp[0][0],-1)
            #print np.shape(teste)
            allarraystemp=[teste2]
        #print np.shape(allarraystemp)
        temptif= None

allarraystemp[allarraystemp==-9999]=np.nan
np.savetxt("allarrays"+".csv", allarraystemp, delimiter=",")


#-----------------------------------------------------#
#-Create tif file to write the correlation values on-#
#-----------------------------------------------------#

print 'Creating new tif files'

#Output name for tif file with calculated correlation
corr_out ="Correlation_Result_NDVI_TEMP_15km_all.tif"
pvalue_out="Correlation_PValue_NDVI_TEMP_15km_all.tif"
corr_sig_out="Correlation_Significant_NDVI_TEMP_15km_all.tif"
print corr_out

# Call the gdal_translate because it needs to be same type of the
float
call(["gdal_translate", "-ot", "Float64","-of", "GTiff","-co",
"TILED=YES", root+"/"+tifname1, root+"/"+corr_out])
call(["gdal_translate", "-ot", "Float64","-of", "GTiff","-co",
"TILED=YES", root+"/"+tifname1, root+"/"+pvalue_out])
call(["gdal_translate", "-ot", "Float64","-of", "GTiff","-co",
"TILED=YES", root+"/"+tifname1, root+"/"+corr_sig_out])

#Open same created file and convert to a array also type float
new_band= gdal.Open(root+"/"+corr_out,gdal.GA_Update)
new_band2= gdal.Open(root+"/"+pvalue_out,gdal.GA_Update)
new_band3= gdal.Open(root+"/"+corr_sig_out,gdal.GA_Update)

array_corr_band=
np.array(new_band.GetRasterBand(1).ReadAsArray(),dtype=float)
array_corr_band2=
np.array(new_band2.GetRasterBand(1).ReadAsArray(),dtype=float)
array_corr_band3=
np.array(new_band3.GetRasterBand(1).ReadAsArray(),dtype=float)
```

```
nodatavalue = new_band.GetRasterBand(1).GetNoDataValue()
print "no data value:", nodatavalue

#-----------------------------------------------#
#Get values from the time series for each pixel#
#-----------------------------------------------#
print "Make pixelwise correlation"
list_corr=[]
list_pvalue=[]

for j in range(len(allarraysndvi[0])):
    #print 'new pixel'
    list_ndvi = []
    list_temp = []
    print 'list lenght', len(list_ndvi)
    for i in range(len(allarraysndvi)):
        pixn = allarraysndvi[i,j]
        pixt = allarraystemp[i,j]
        list_ndvi.append(pixn)
        list_temp.append(pixt)
    #print 'ndvi len'
    #print len(list_ndvi)
    x = np.asarray(list_ndvi)
    y = np.asarray(list_temp)
    #idx1 = np.logical_and(x != nodatavalue , y != nodatavalue)
    #idx1=np.logical_and(np.logical_not(np.isnan(x)) ,
np.logical_not(np.isnan(y)))
    #x=x[idx1]
    #y=y[idx1]
    pixcorr = pearsonr(x,y)
    #print pixcorr[0]
    list_corr.append(pixcorr[0])
    list_pvalue.append(pixcorr[1])

print 'list corr len'
print len(list_corr)
#Saves data in csv file
np.savetxt("list_corr.csv", list_corr, delimiter=",")
np.savetxt("list_pvalue.csv", list_pvalue, delimiter=",")

#------ Reshaping and writing new files-------#

array_corr_band = np.asarray(list_corr)
array_corr_band2 = np.asarray(list_pvalue)

array_corr_band3 = np.asarray(list_corr)
# creating an array with pixels with a significance over 0.05 (p-
value)
array_corr_band3[array_corr_band2>0.05] = nodatavalue

print np.shape(array_corr_band)
tif = gdal.Open(root+"/"+tifname1,gdal.GA_Update)
cols = tif.RasterXSize
rows = tif.RasterYSize
print cols, rows
array_corr_band=np.reshape(array_corr_band,(rows,cols))
array_corr_band2=np.reshape(array_corr_band2,(rows,cols))
array_corr_band3=np.reshape(array_corr_band3,(rows,cols))
print np.shape(array_corr_band)
```

```
print 'Writing new data'
# Write array write data on the new tiffs
new_band.GetRasterBand(1).WriteArray(array_corr_band)
new_band2.GetRasterBand(1).WriteArray(array_corr_band2)
new_band3.GetRasterBand(1).WriteArray(array_corr_band3)

new_band = None
new_band2 = None
new_band3 = None

#-----------------------------------------------------#

print 'Done'
```

Appendix A.10 Save data to csv file

```
# -*- coding: utf-8 -*-
"""
@author: JFG
"""
from osgeo import gdal
from osgeo.gdalconst import *
import os,fnmatch
import numpy as np
import sys
import csv
from subprocess import call


#folders
inputFolder="C:/Users/Joana/Documents/Tese_Computador/Datasets_Originais_TESE/NDVI_O/MarOut_NDVI/1_Marco"

#pattern of the tif files
pattern_tif='AVERAGE_MARCH'+'*'+'.[Tt][Ii][Ff]'

L1 =  list()
for root, dirs, files in os.walk(inputFolder):

    for tifname in fnmatch.filter(files,pattern_tif):


        a = tifname.find(".tif")
        prefix = tifname[:a-3]

        filename = gdal.Open(root+"/"+tifname,gdal.GA_Update)
        print tifname
        tif1 =
[[np.array(filename.GetRasterBand(1).ReadAsArray(),dtype=float)]]

        #Gets the no data value for each file
        nodatavaluetif = filename.GetRasterBand(1).GetNoDataValue()

        filename = None

        n = np.asarray(tif1)

        print 'SHAPES'
        print np.shape(n)

        n = np.reshape(n[0][0],-1)

        print 'no data:'
        print nodatavaluetif

        n = n[np.logical_not(np.isnan(n))]
        n = n[n != nodatavaluetif]
        print np.shape(n)
        m = np.average(n)
        L1.append(m)

np.savetxt(prefix+".csv", L1, delimiter=",")
print 'Done!'
```

Appendix B Monthly NDVI anomalies

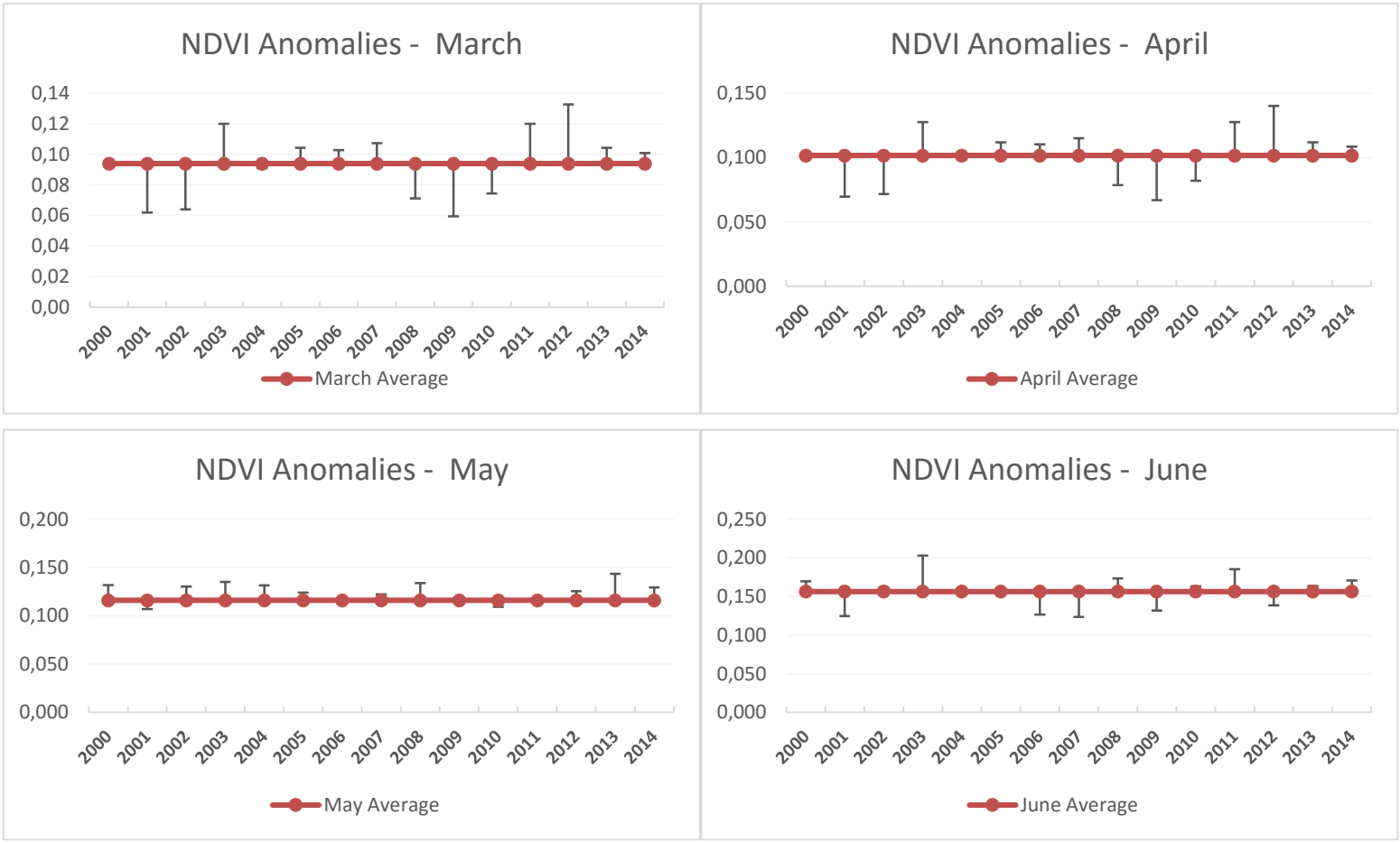


Figure 7.1 | NDVI Anomalies for March, April, May and June, with reference to the respective monthly average for the period 2000-2014.

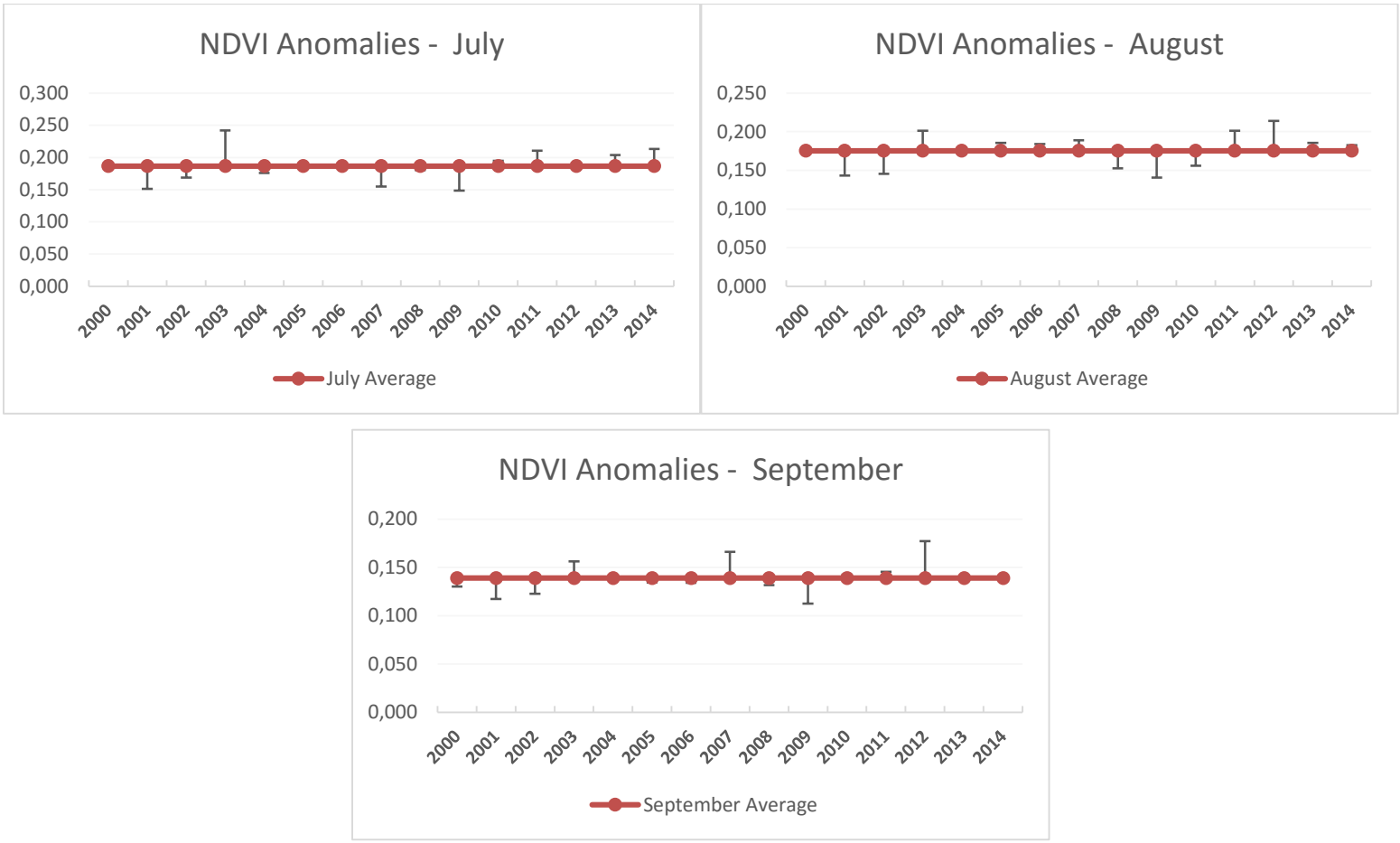


Figure 7.2 | NDVI Anomalies for July, August and September, with reference to the respective monthly average for the period 2000-2014.

Appendix C – Spatial Distribution of Monthly NDVI anomalies

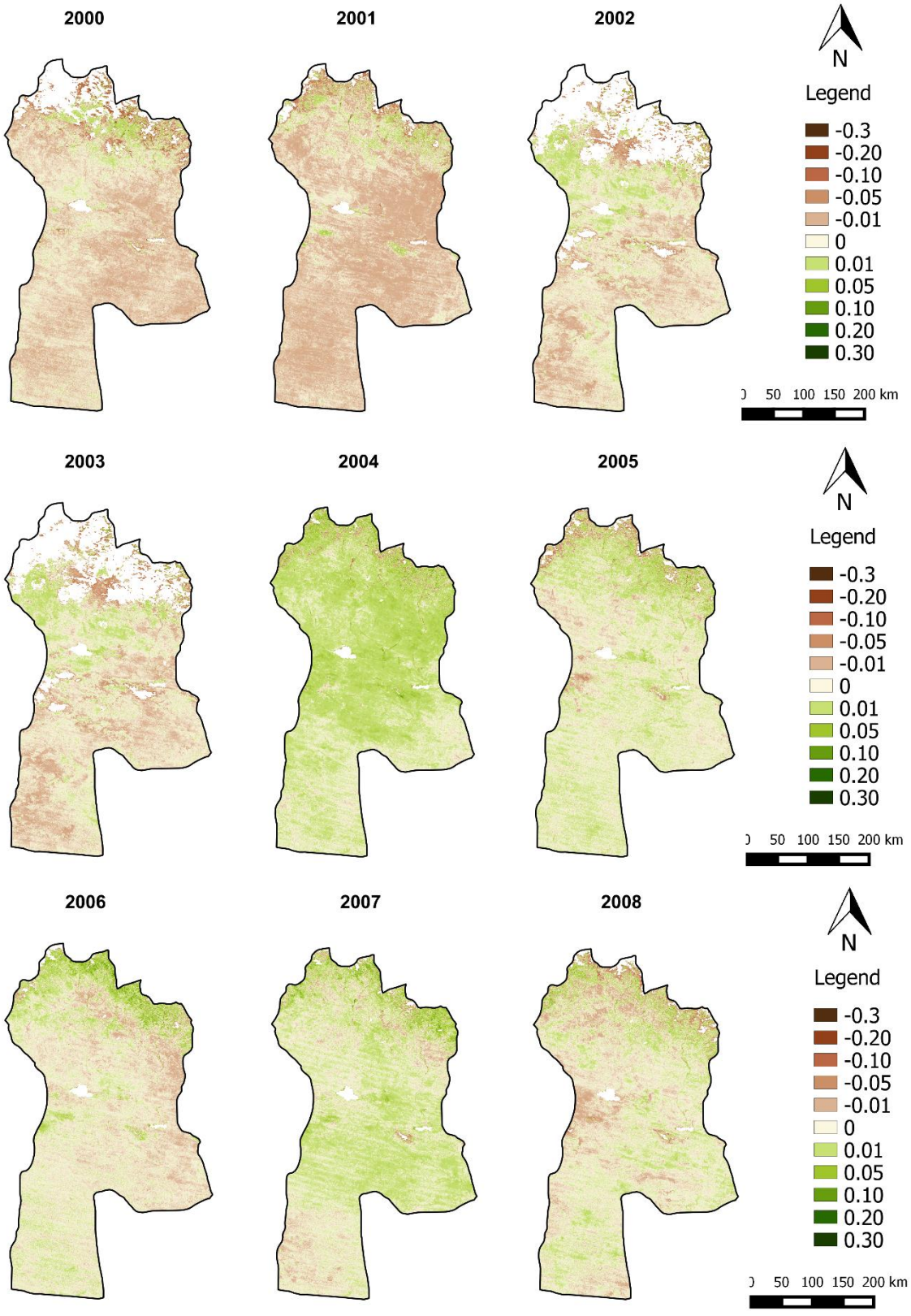


Figure 7.3 | Spatial distribution of NDVI anomalies of March for the period 2000 – 2008.

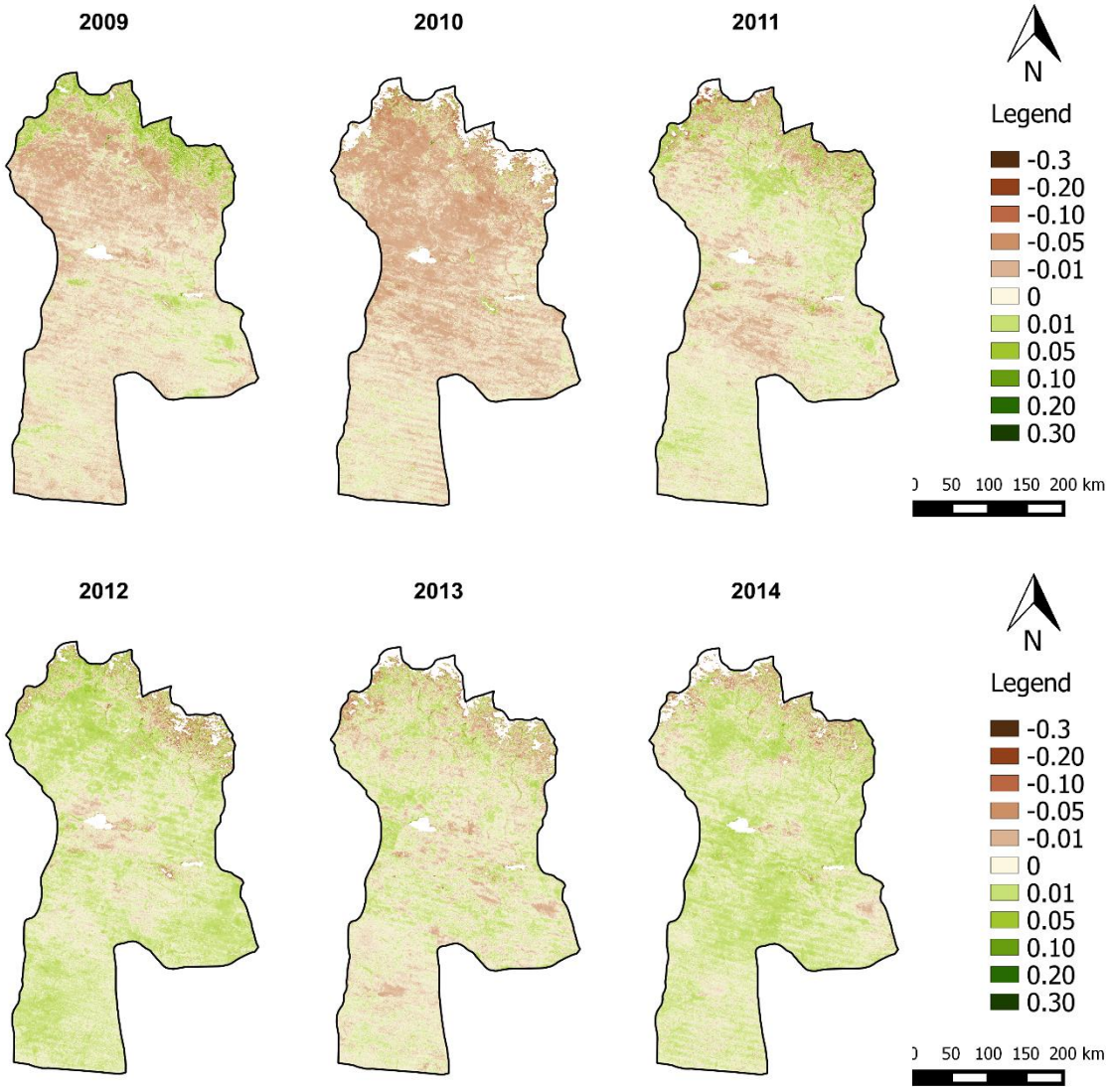


Figure 7.4 | Spatial distribution of NDVI anomalies of March for the period 2009 – 2014.

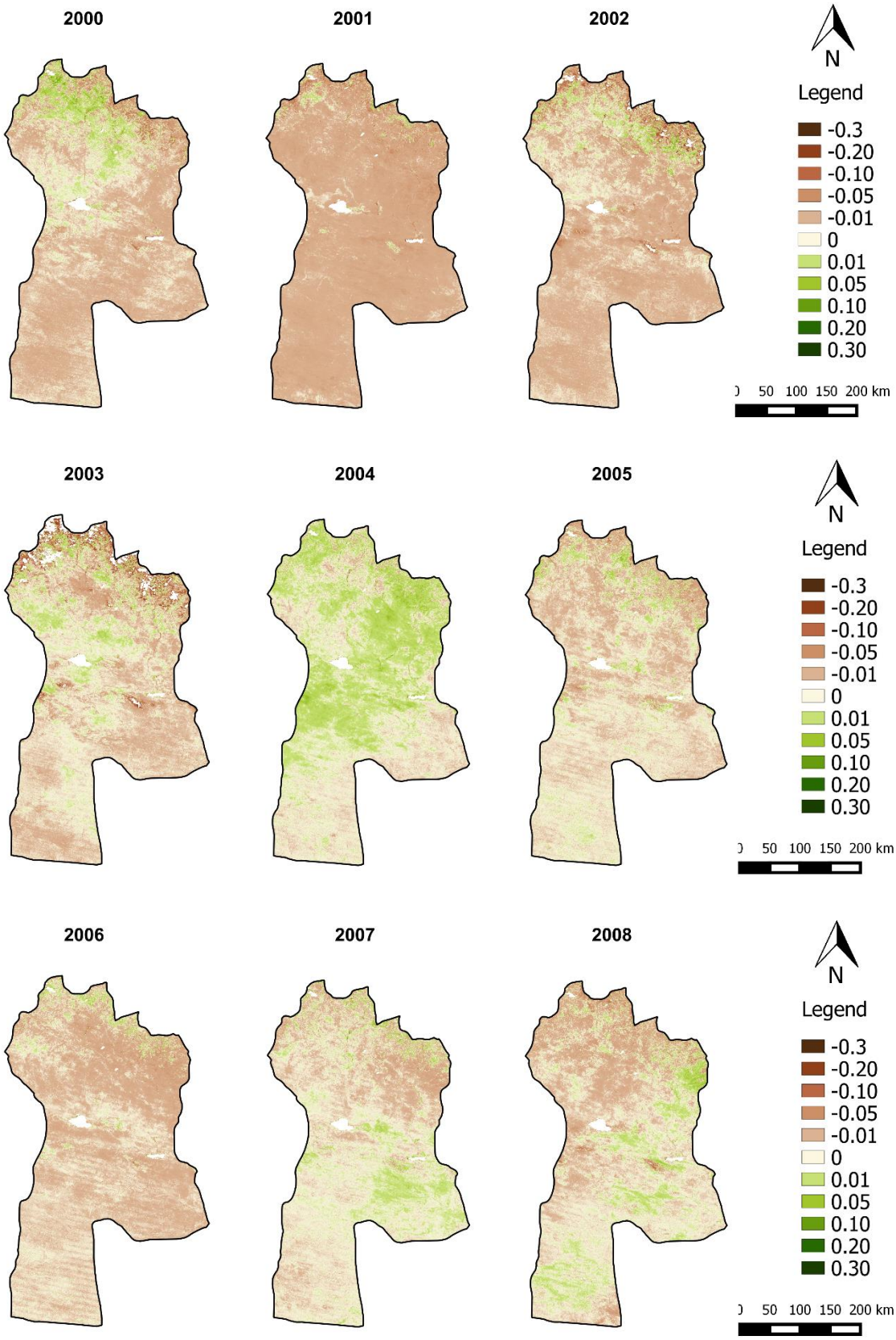


Figure 7.5 | Spatial distribution of NDVI anomalies of April for the period 2000 – 2008.

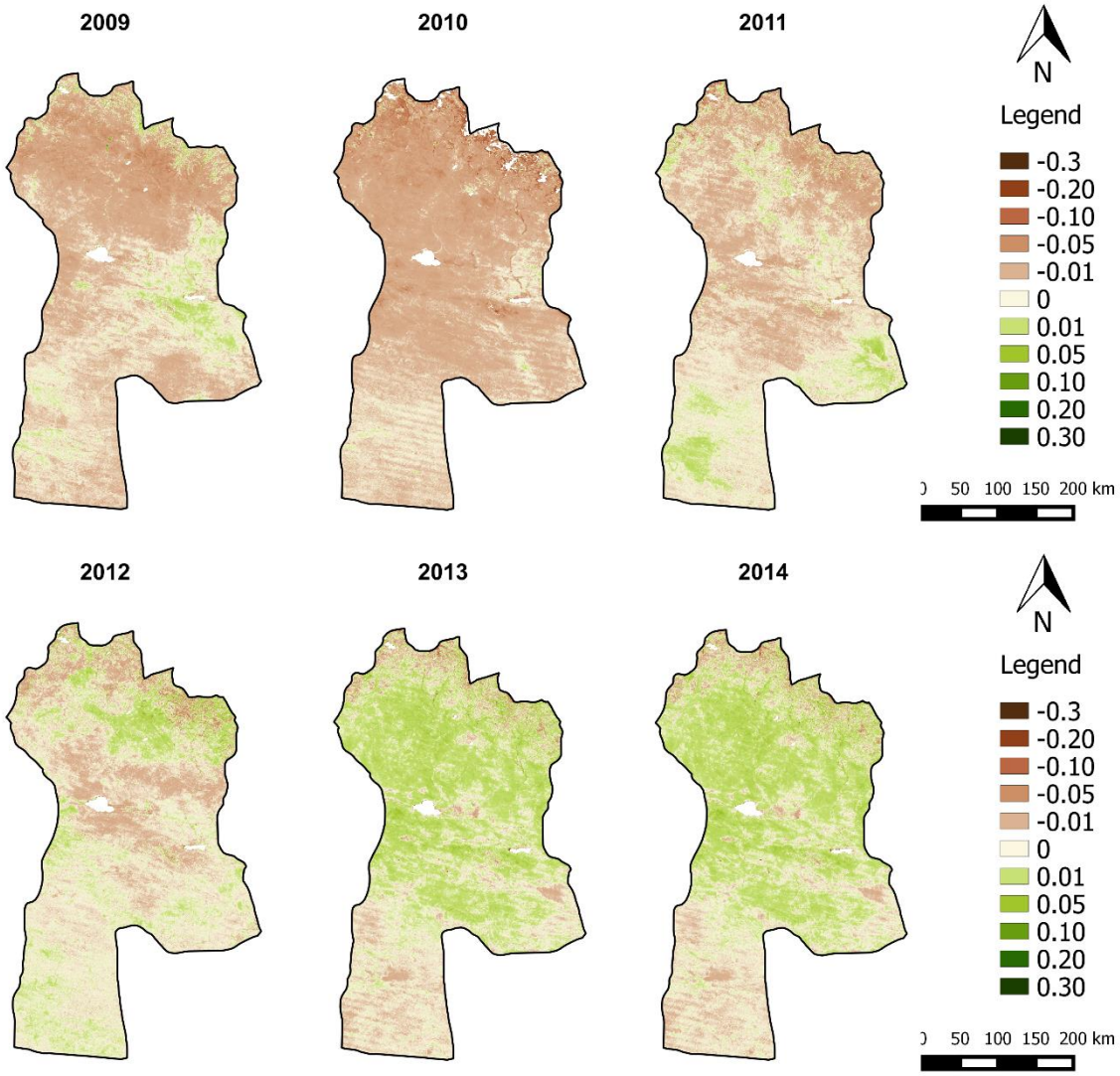


Figure 7.6 | Spatial distribution of NDVI anomalies of April for the period 2009 – 2014.

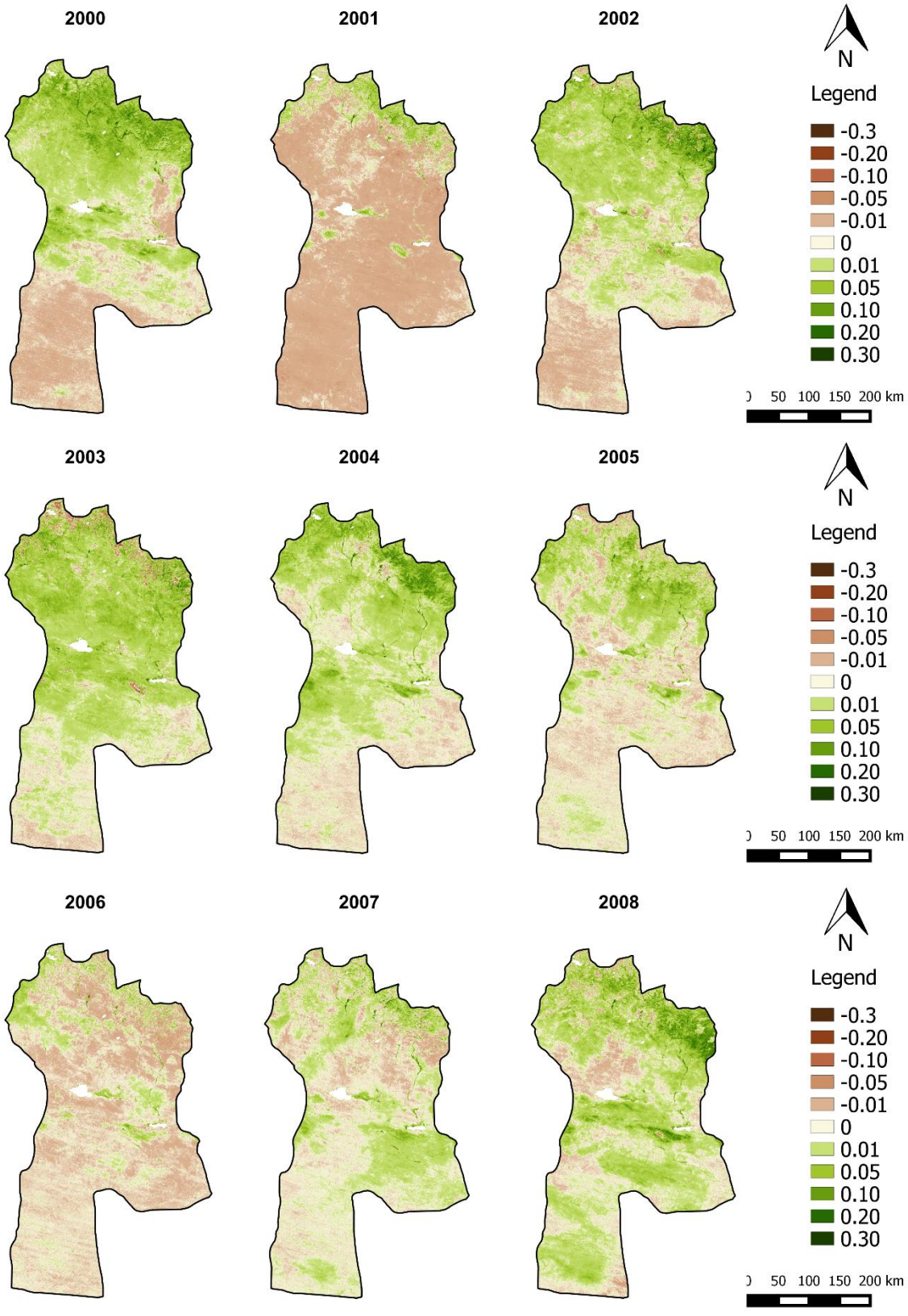


Figure 7.7 | Spatial distribution of NDVI anomalies of May for the period 2000 – 2008.

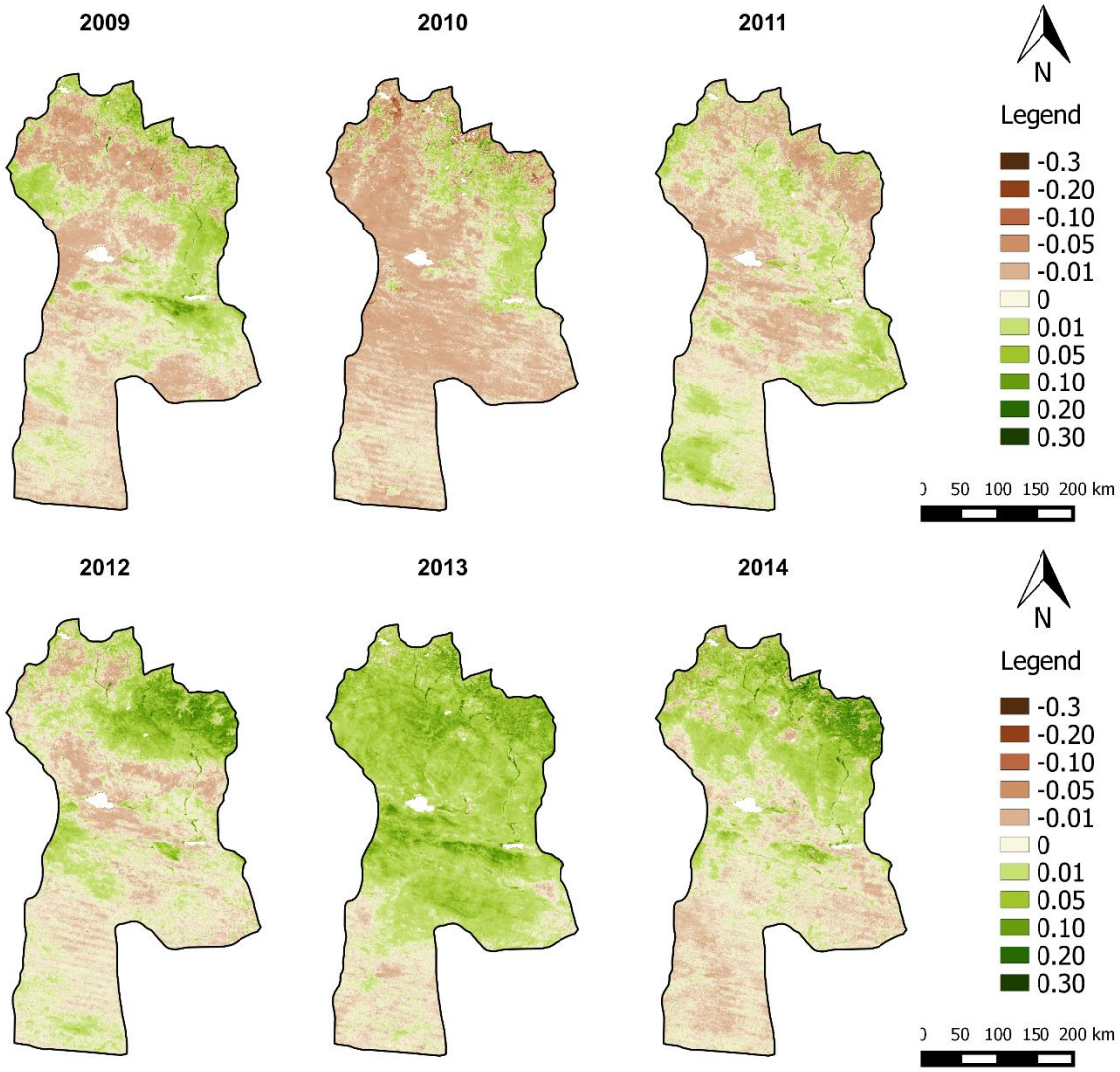


Figure 7.8 | Spatial distribution of NDVI anomalies of May for the period 2009 – 2014.

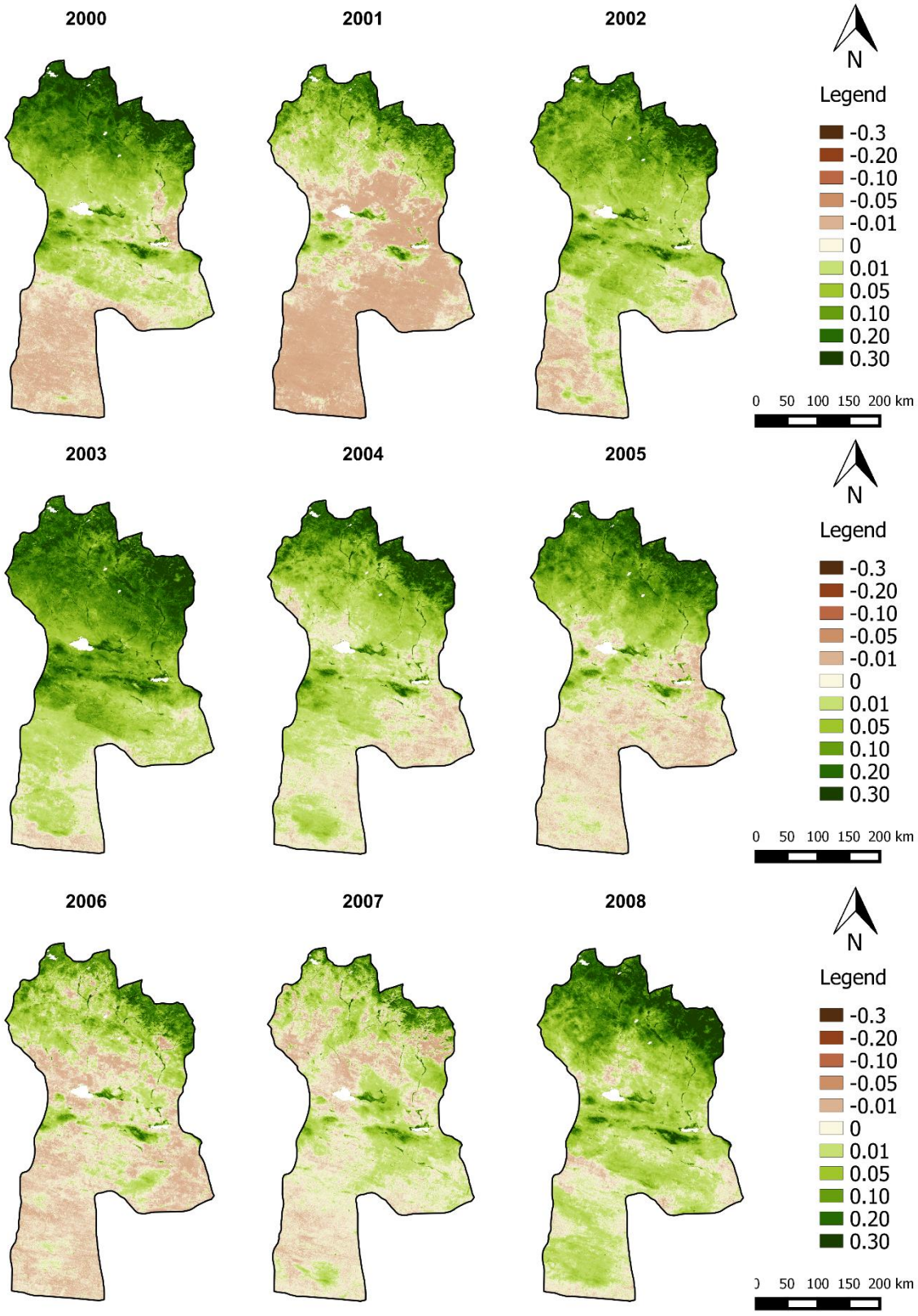


Figure 7.9 | Spatial distribution of NDVI anomalies of June for the period 2000 – 2008.

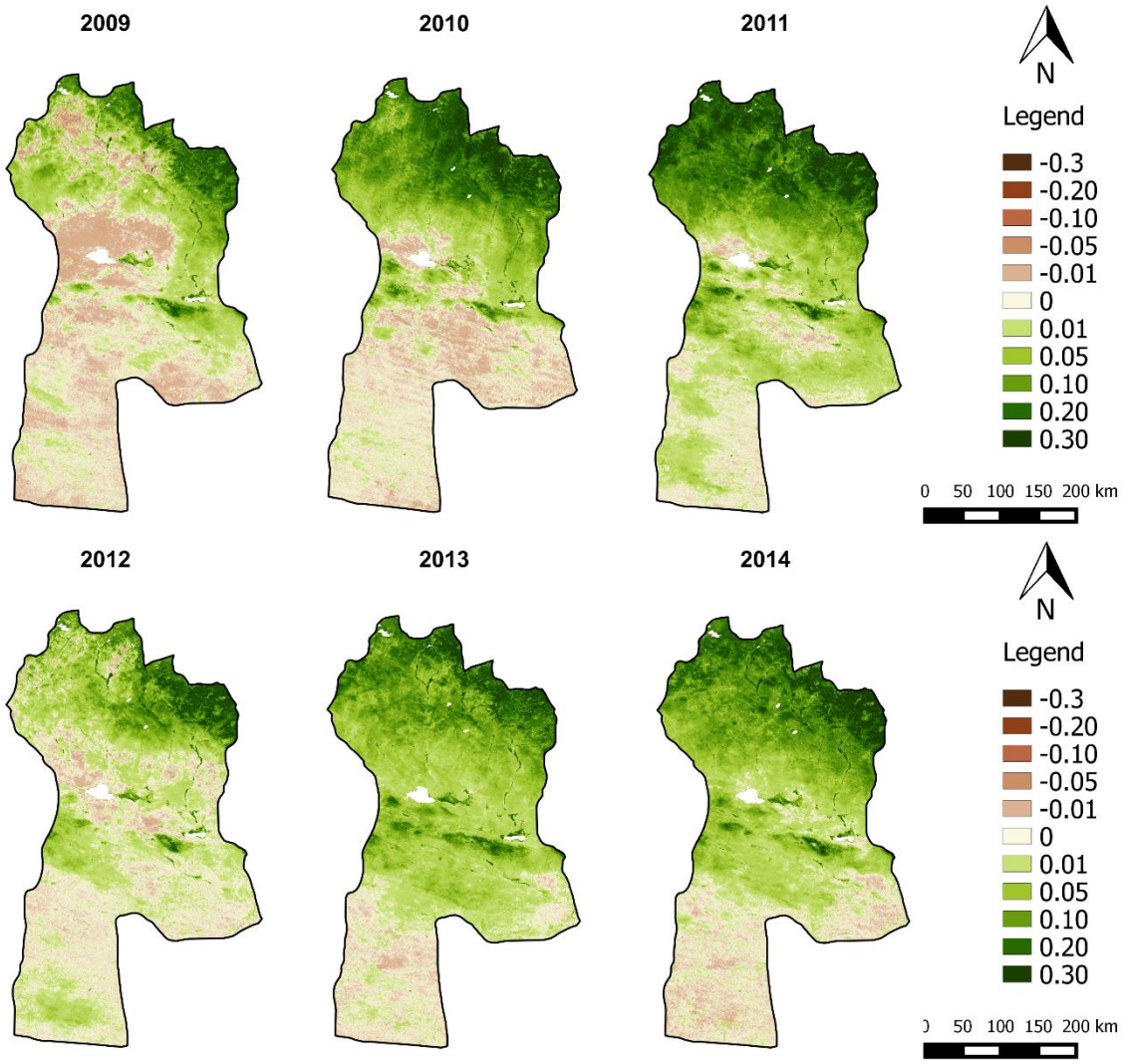


Figure 7.10 | Spatial distribution of NDVI anomalies of June for the period 2009 – 2014.

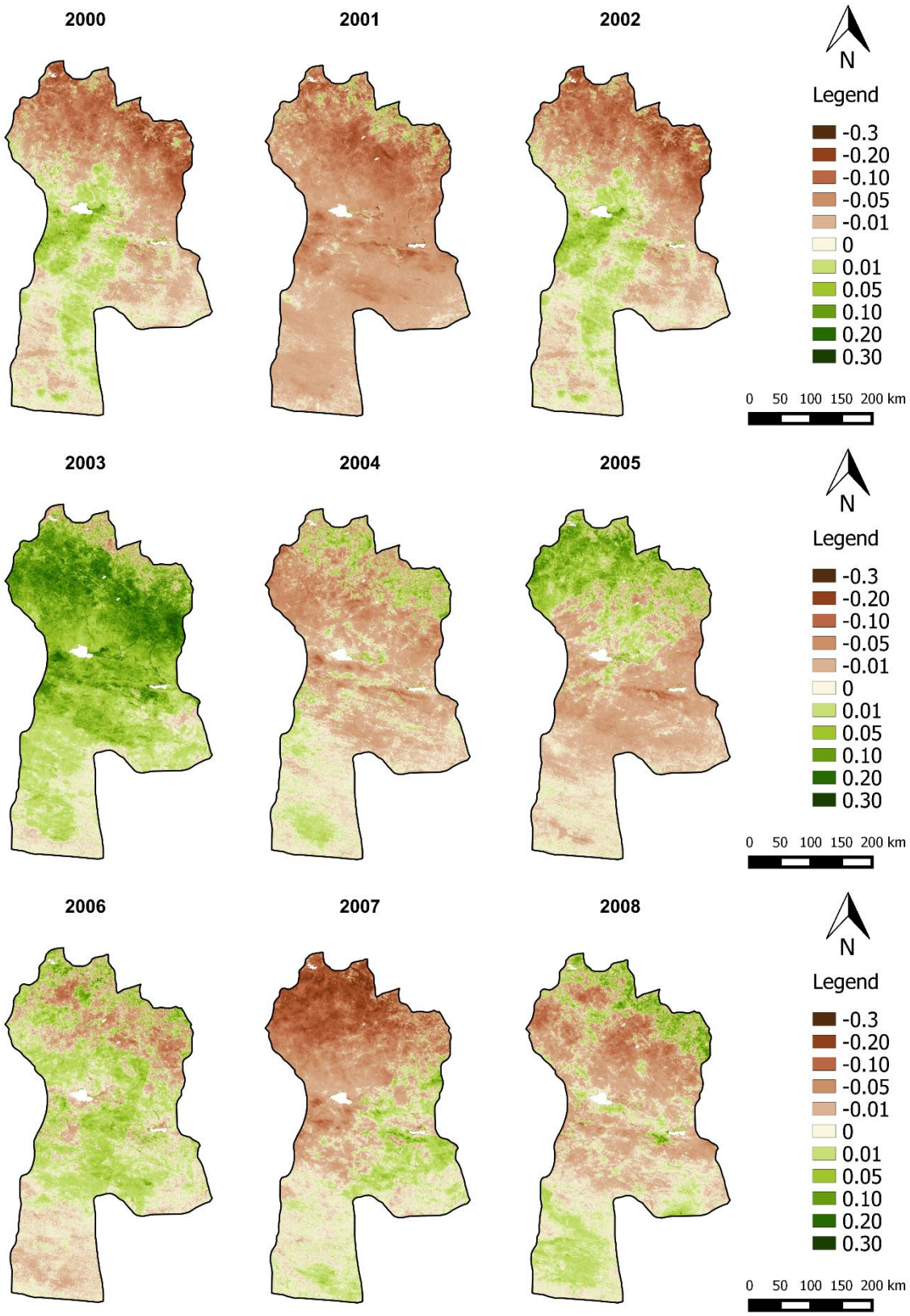


Figure 7.11 | Spatial distribution of NDVI anomalies of July for the period 2000 – 2008.

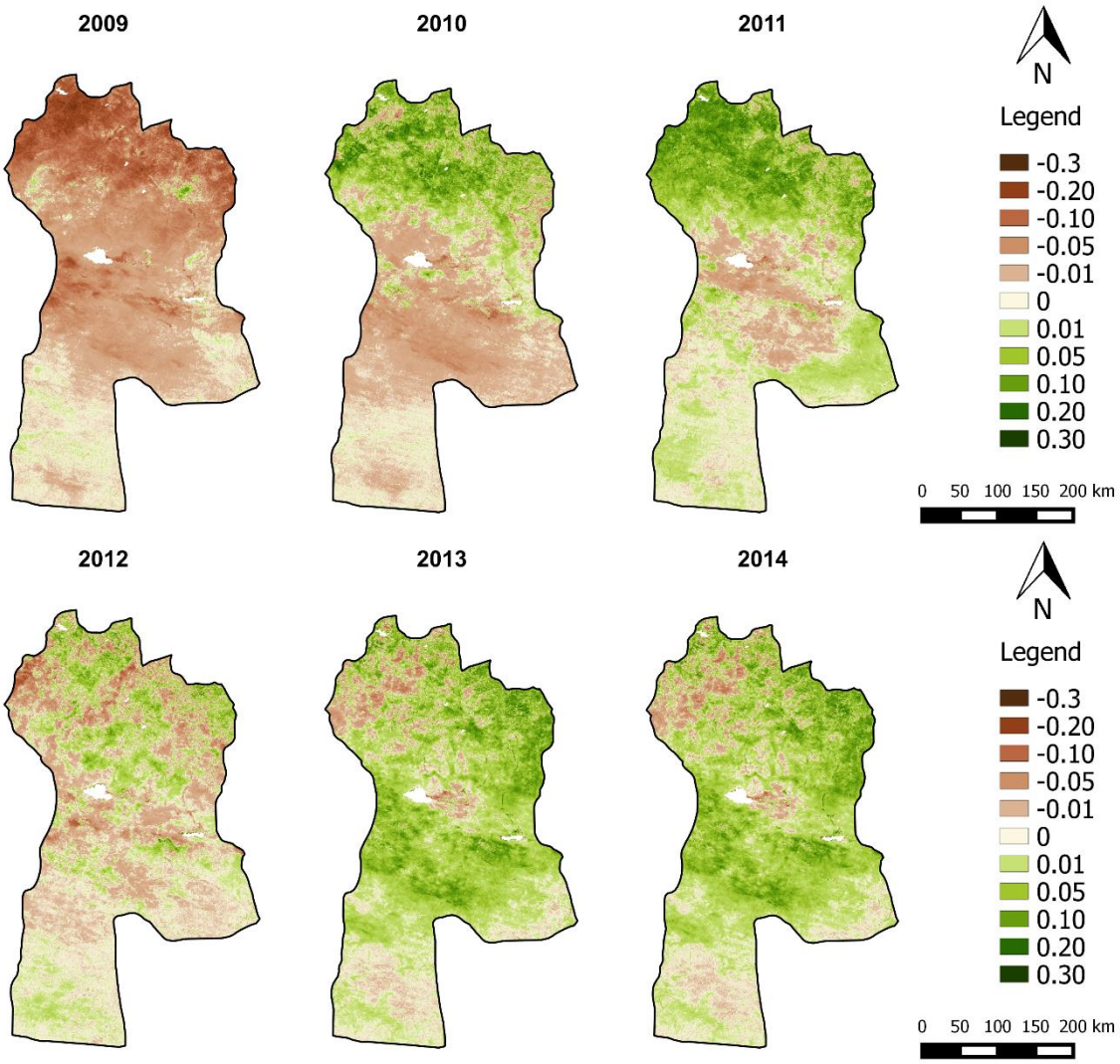


Figure 7.12 | Spatial distribution of NDVI anomalies of July for the period 2009 – 2014.

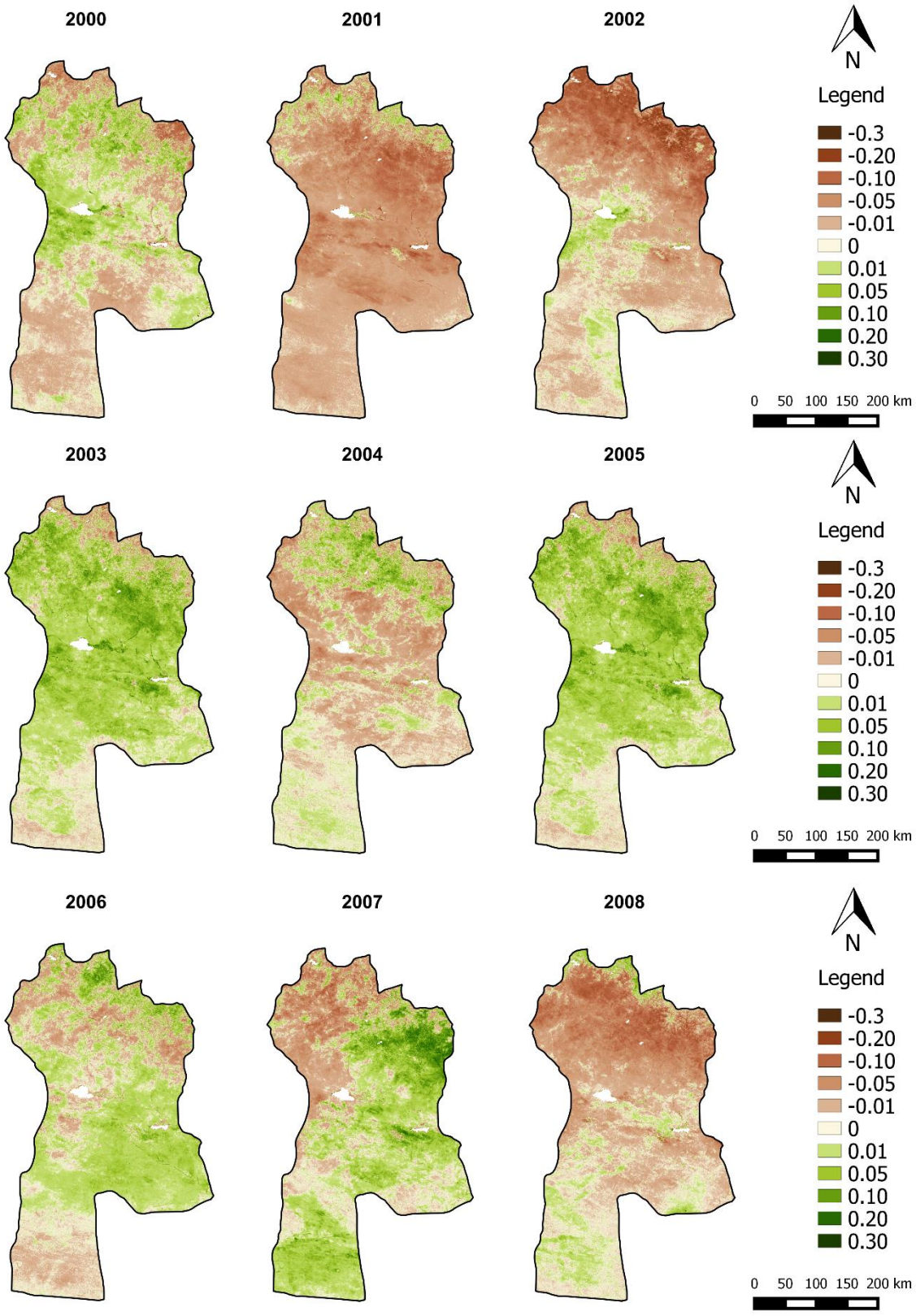


Figure 7.13 | Spatial distribution of NDVI anomalies of August for the period 2000 – 2008.

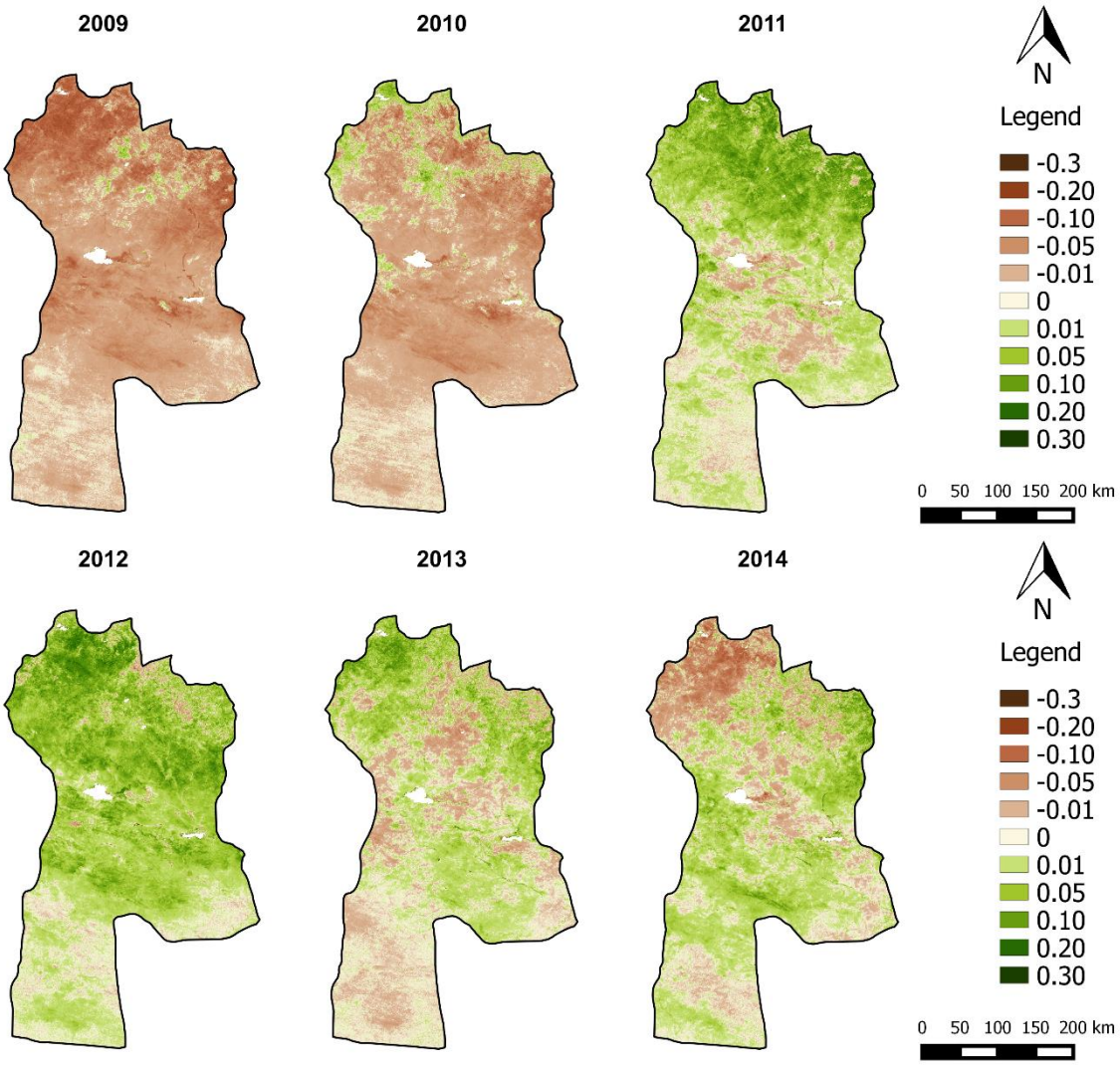


Figure 7.14 | Spatial distribution of NDVI anomalies of August for the period 2009 – 2014.

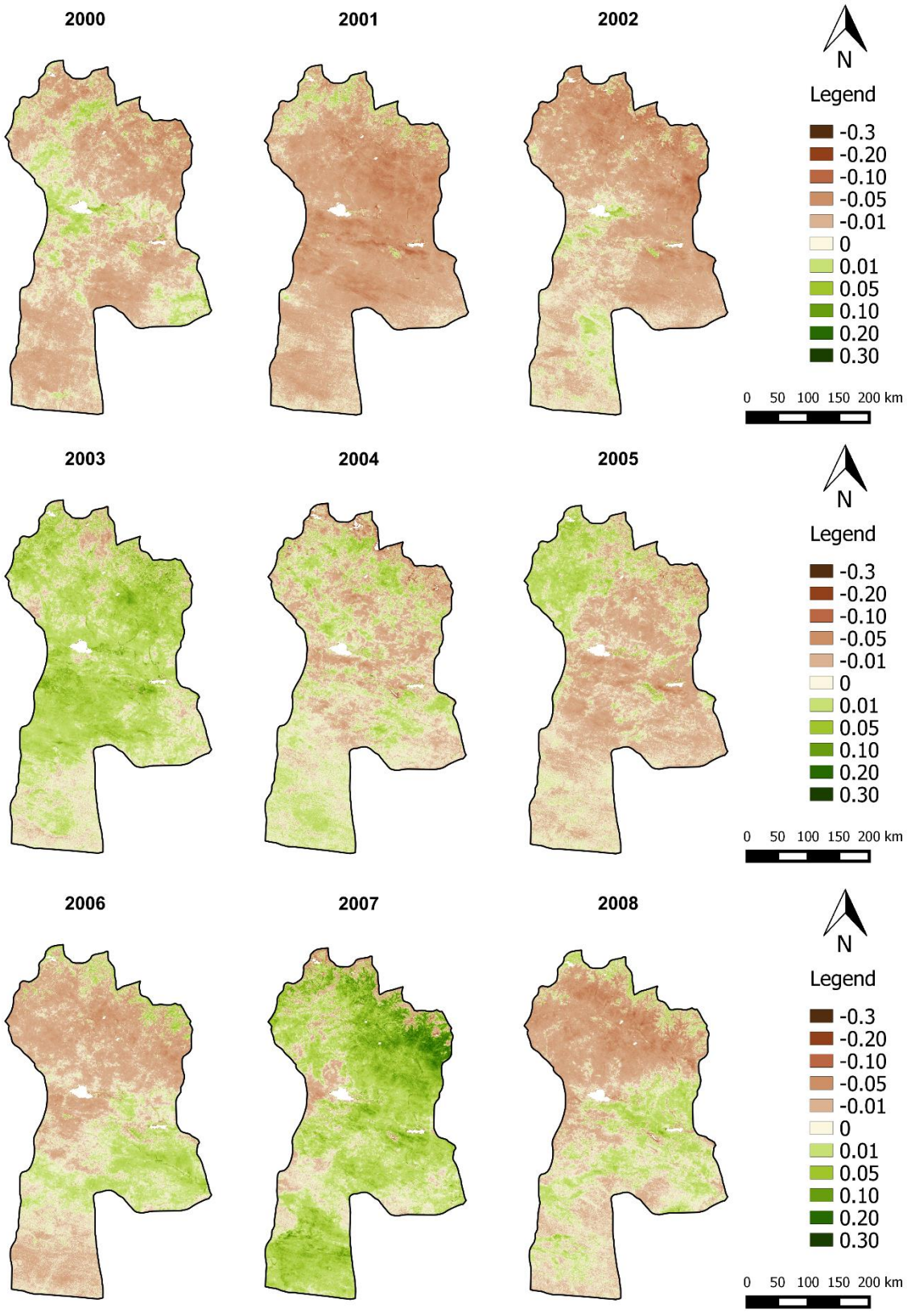


Figure 7.15 | Spatial distribution of NDVI anomalies of September for the period 2000 – 2008.

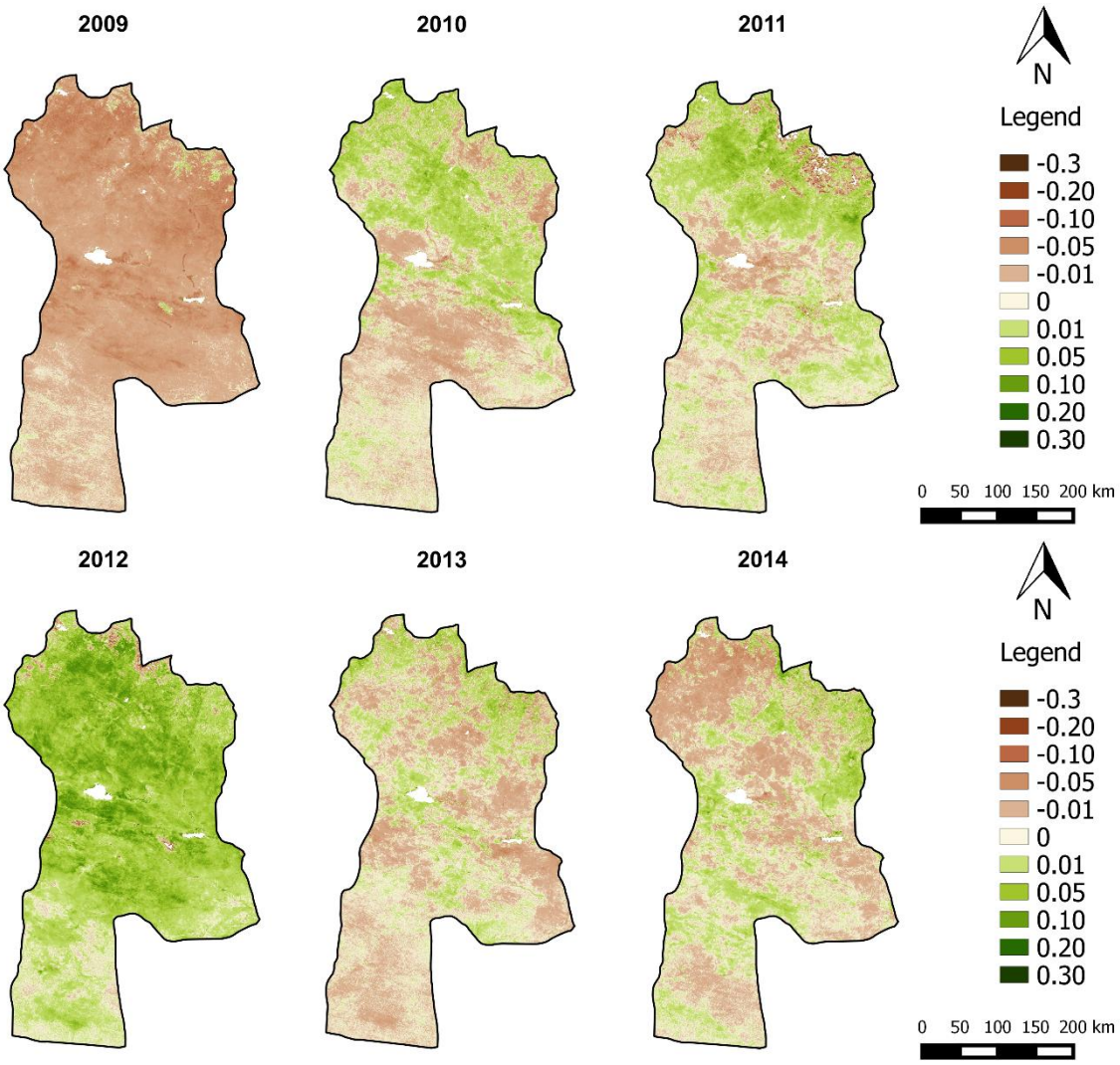


Figure 7.16 | Spatial distribution of NDVI anomalies of September for the period 2009 – 2014.

Appendix D – Monthly Land Surface Temperature anomalies

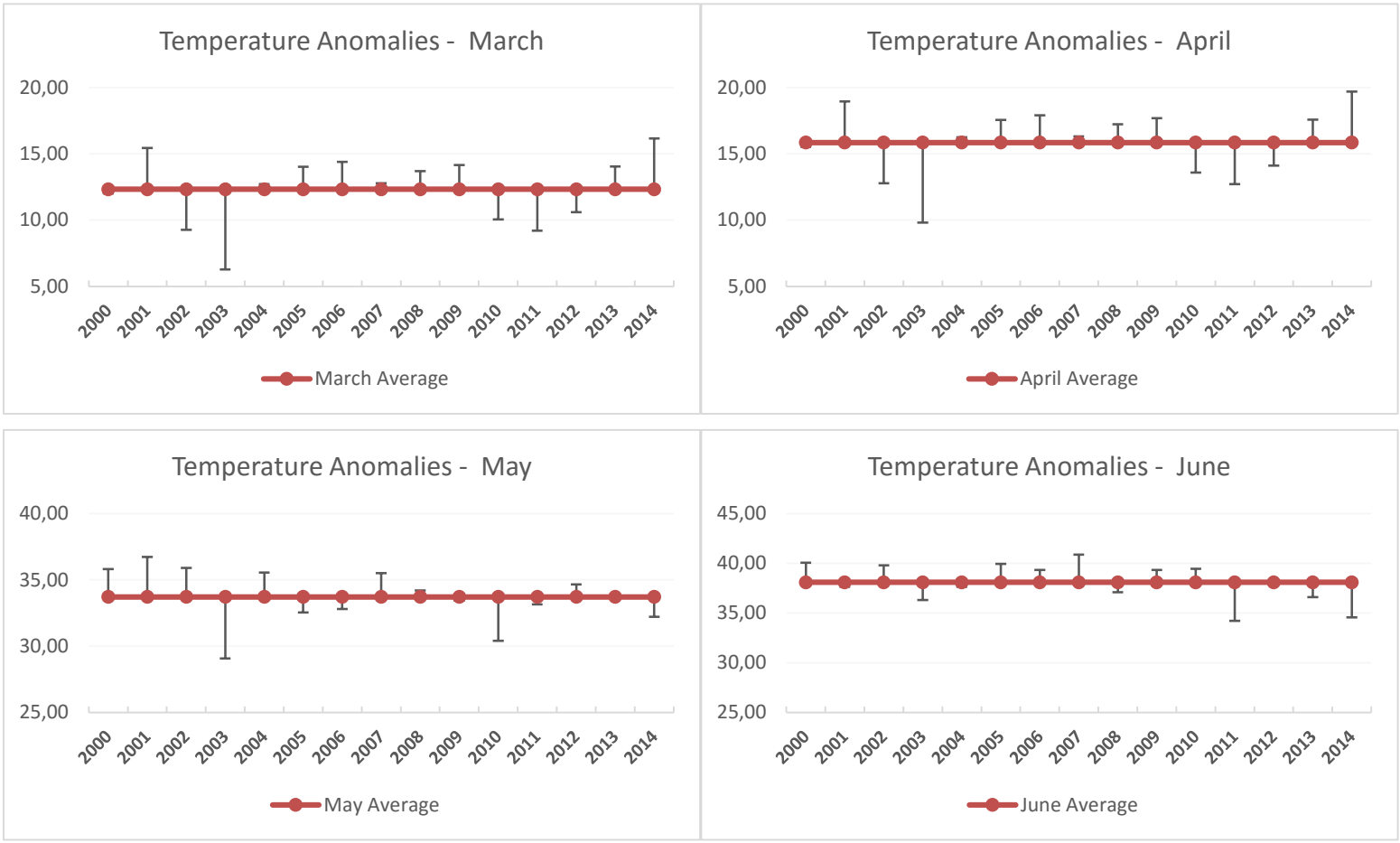


Figure 7.17 | Land Surface Temperature anomalies for the time series 2000-2014 (March – June) (ºC)

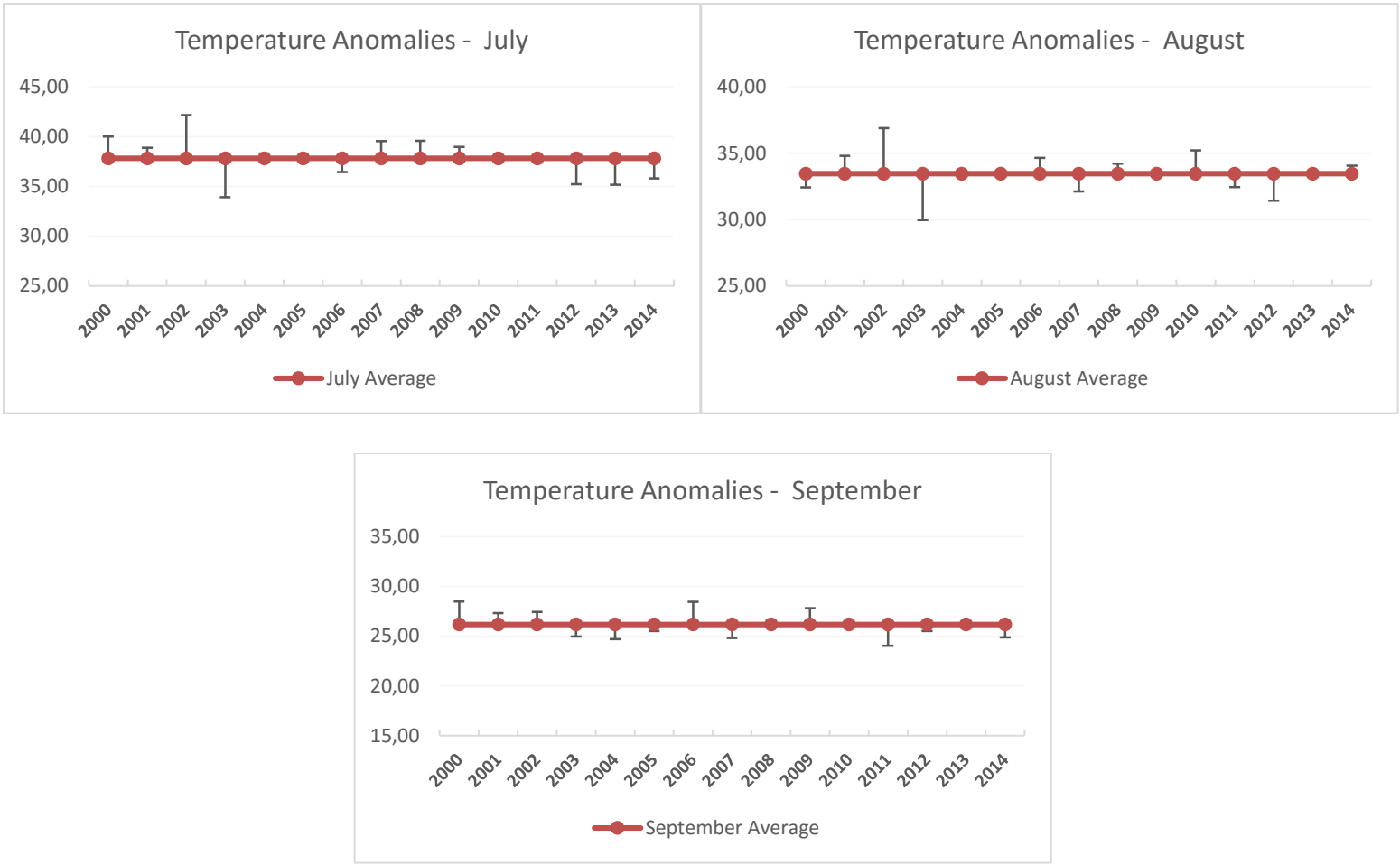


Figure 7.18 | Land Surface Temperature anomalies for the time series 2000-2014 (July – September) (ºC)

Appendix E – Spatial Distribution of Land Surface Temperature anomalies

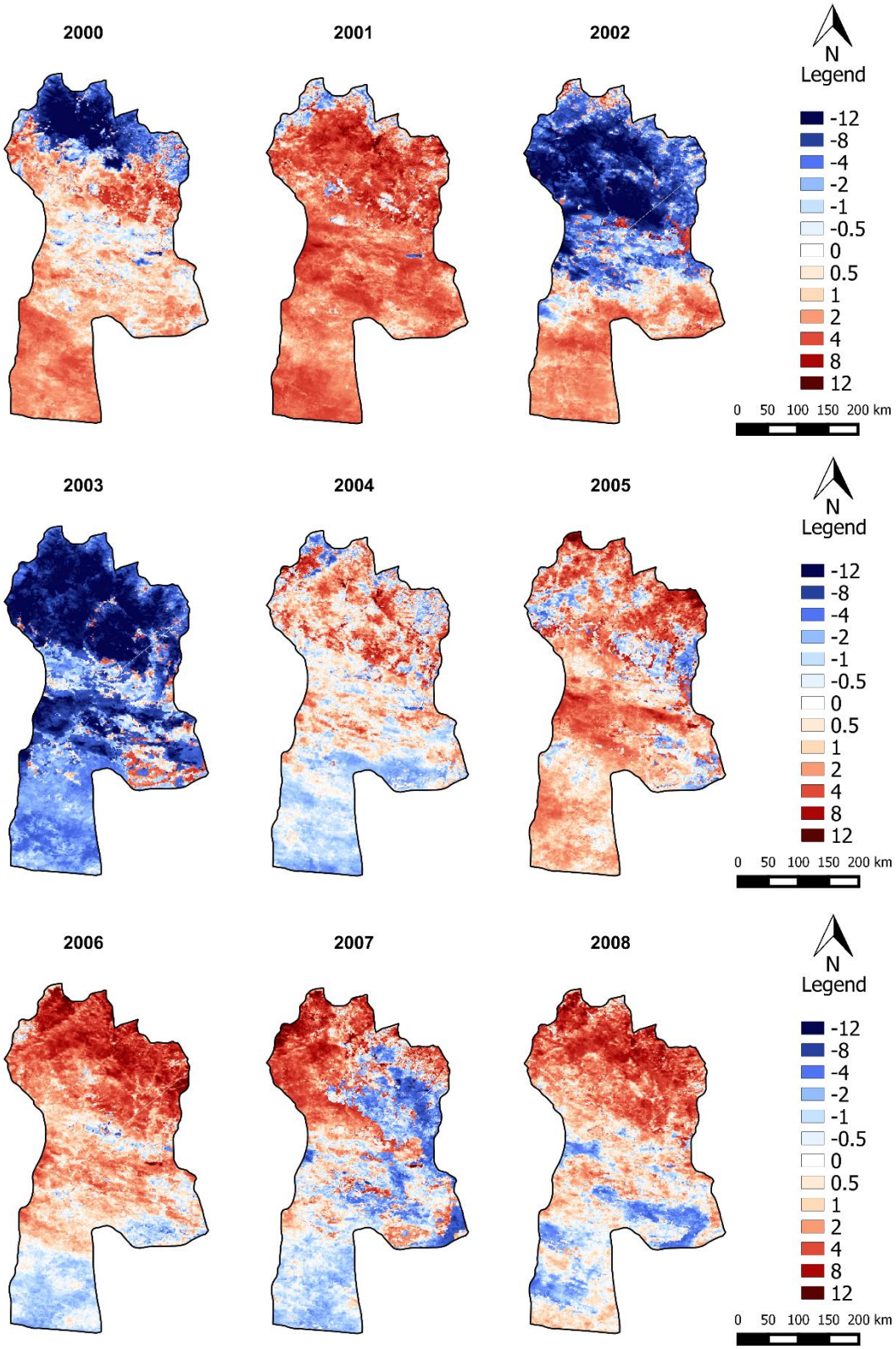


Figure 7.19 | Spatial distribution of LST anomalies of March for the period 2000 – 2008.

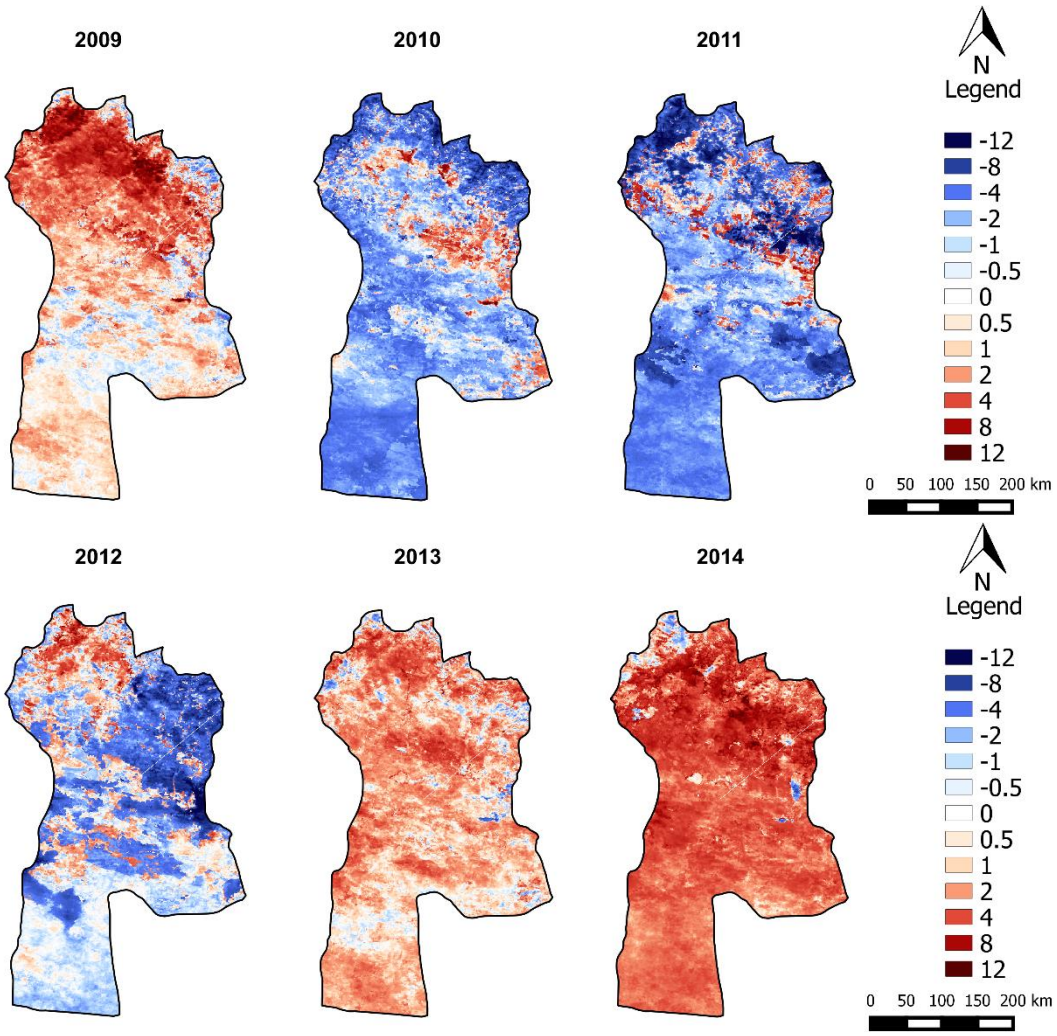


Figure 7.20 | Spatial distribution of LST anomalies of March for the period 2009 – 2014.

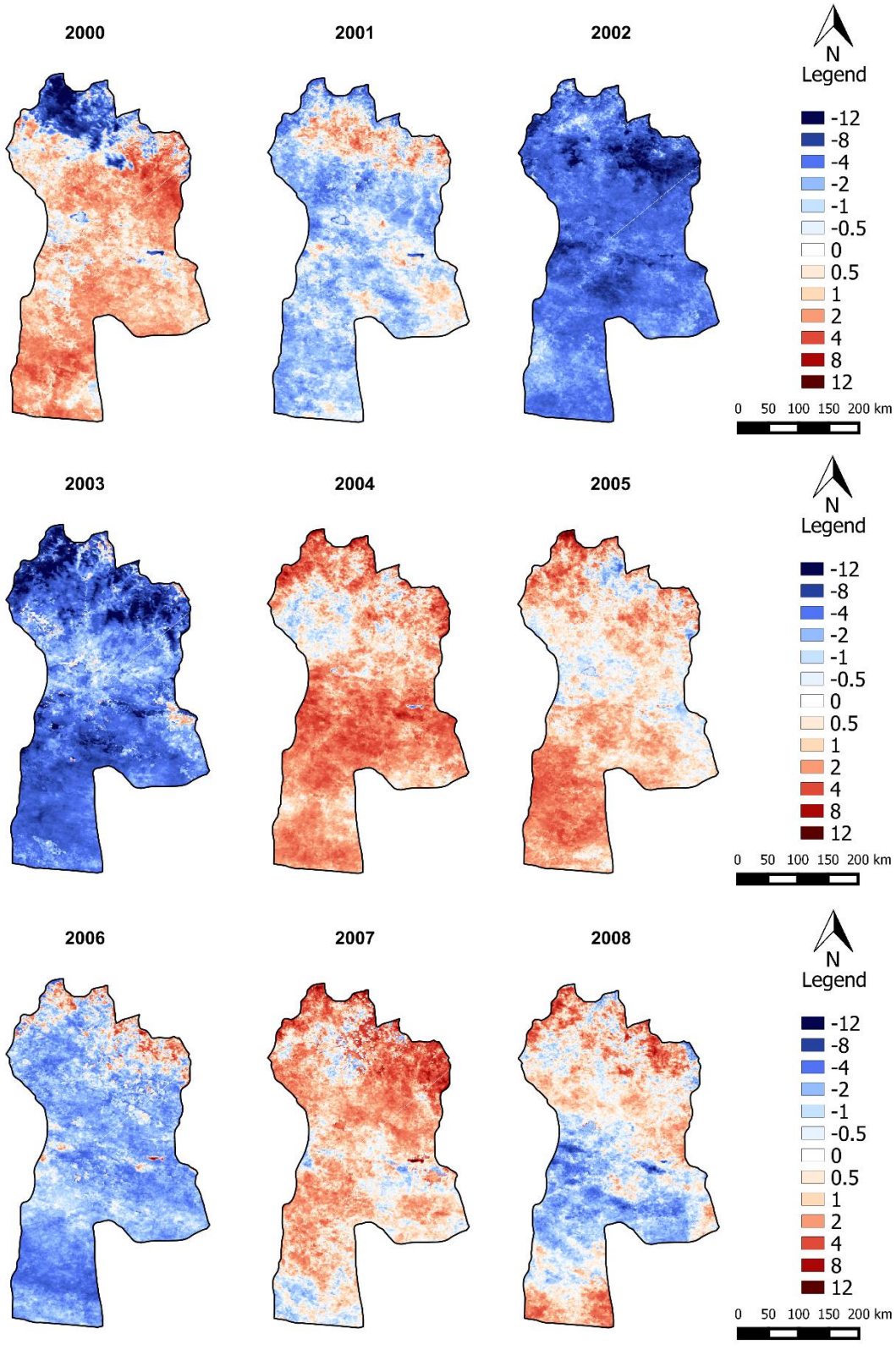


Figure 7.21 | Spatial distribution of LST anomalies of April for the period 2000 – 2008.

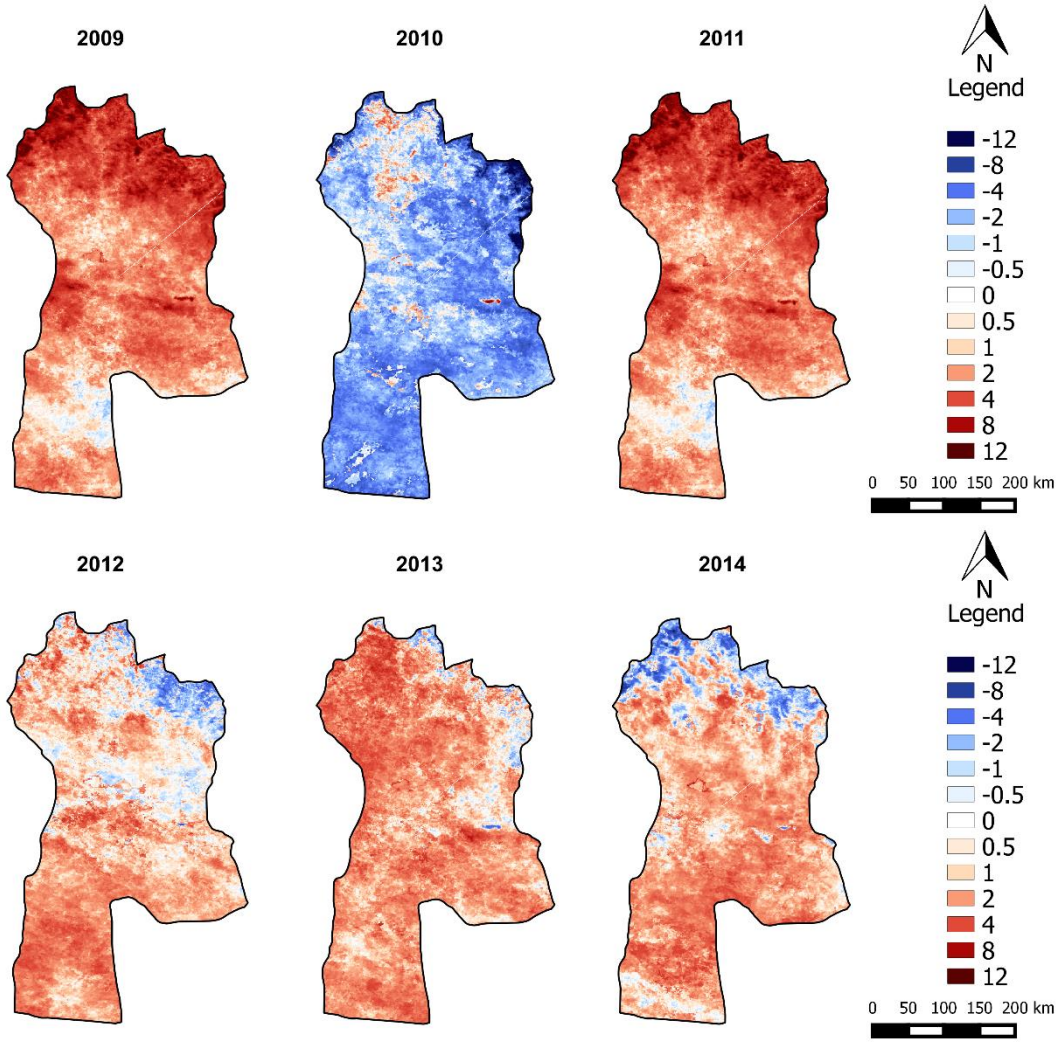


Figure 7.22 | Spatial distribution of LST anomalies of April for the period 2009 – 2014.

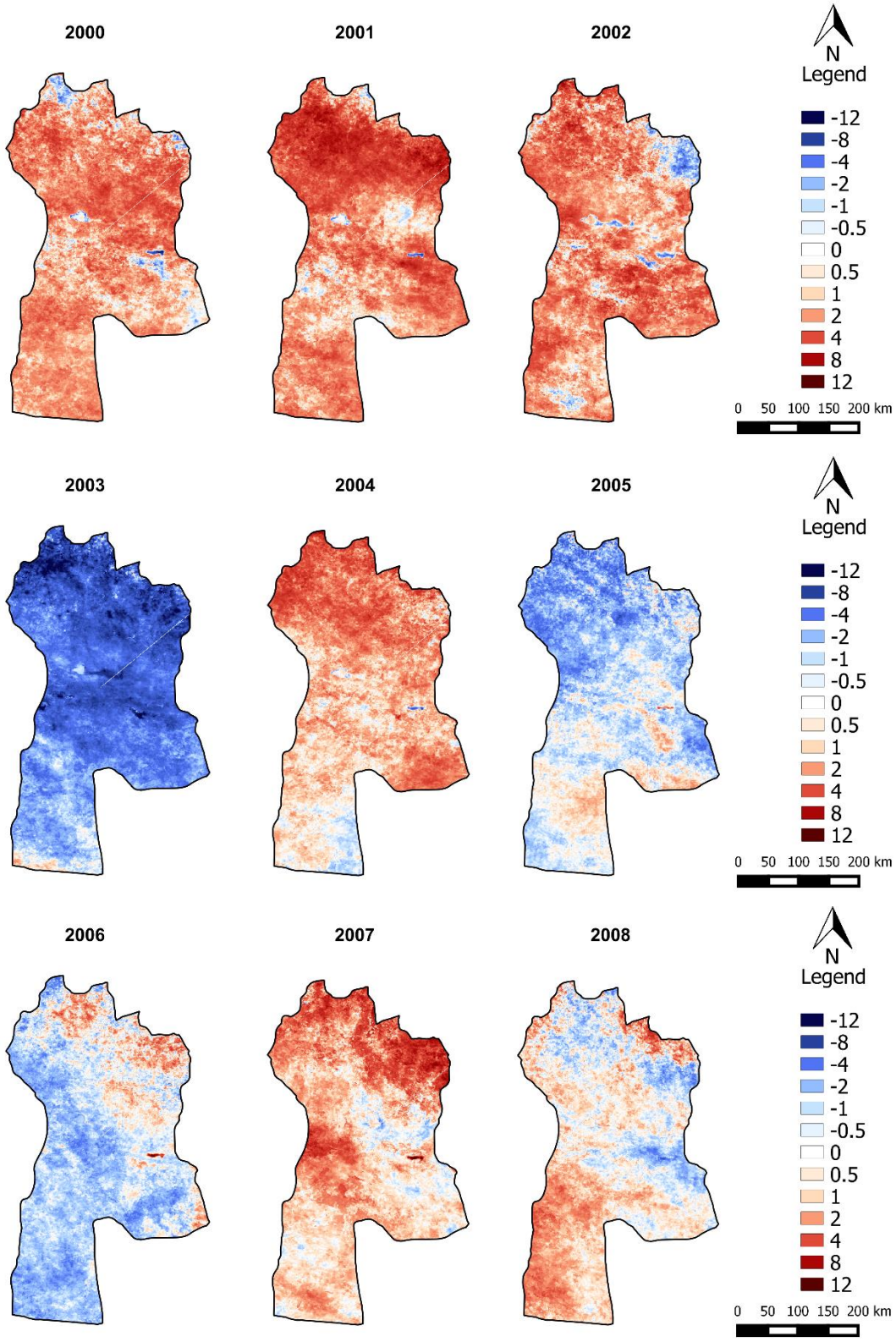


Figure 7.23 | Spatial distribution of LST anomalies of May for the period 2000 – 2008.

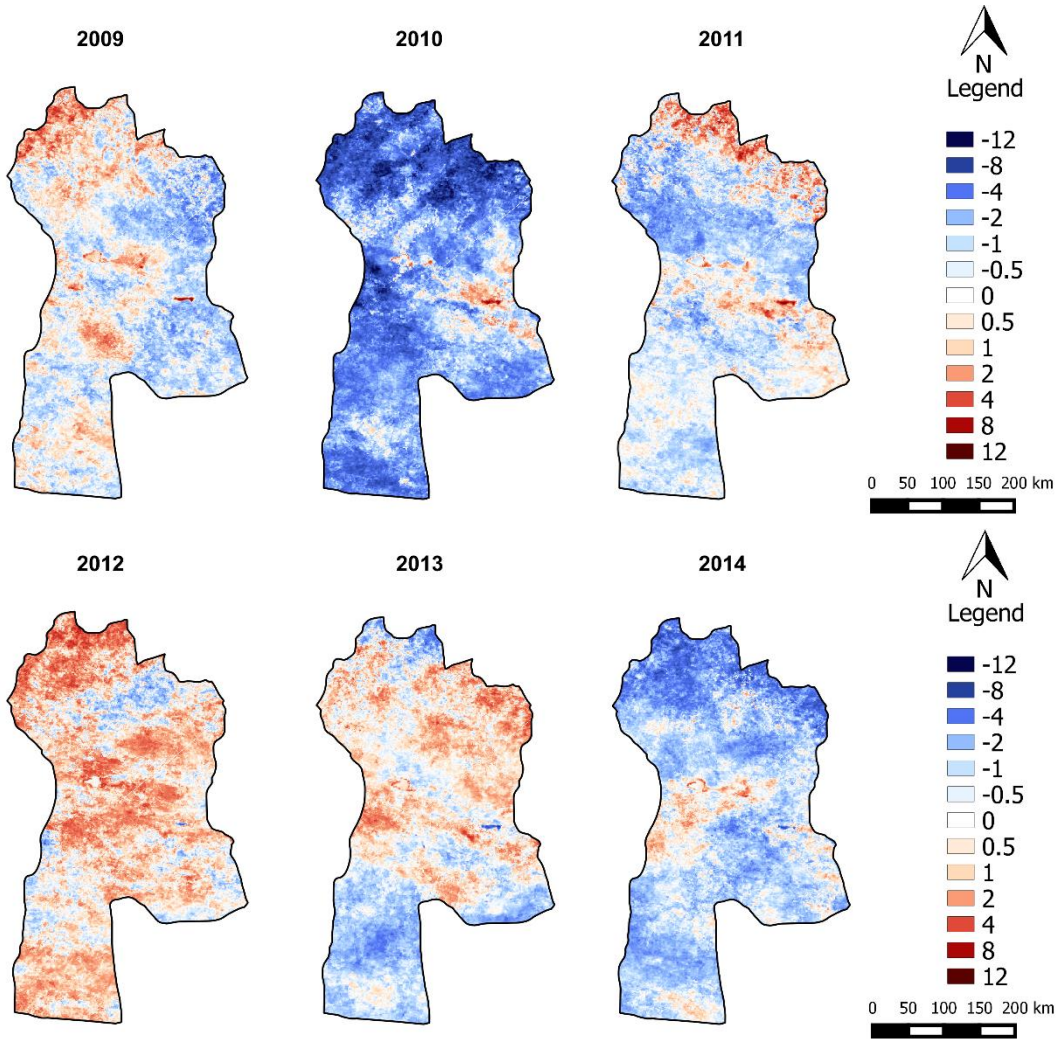


Figure 7.24 | Spatial distribution of LST anomalies of May for the period 2009 – 2014.

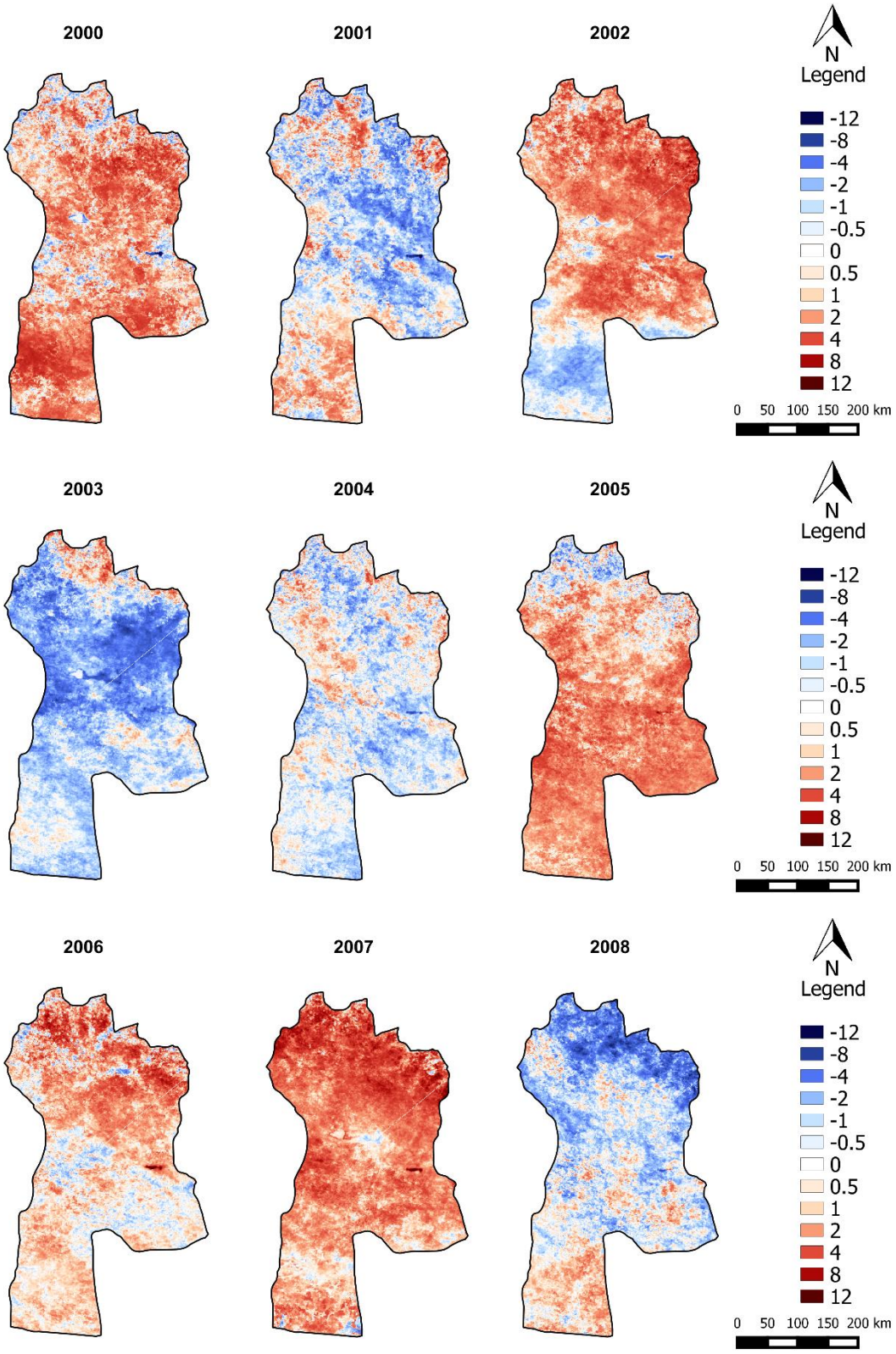


Figure 7.25 | Spatial distribution of LST anomalies of June for the period 2000 – 2008.

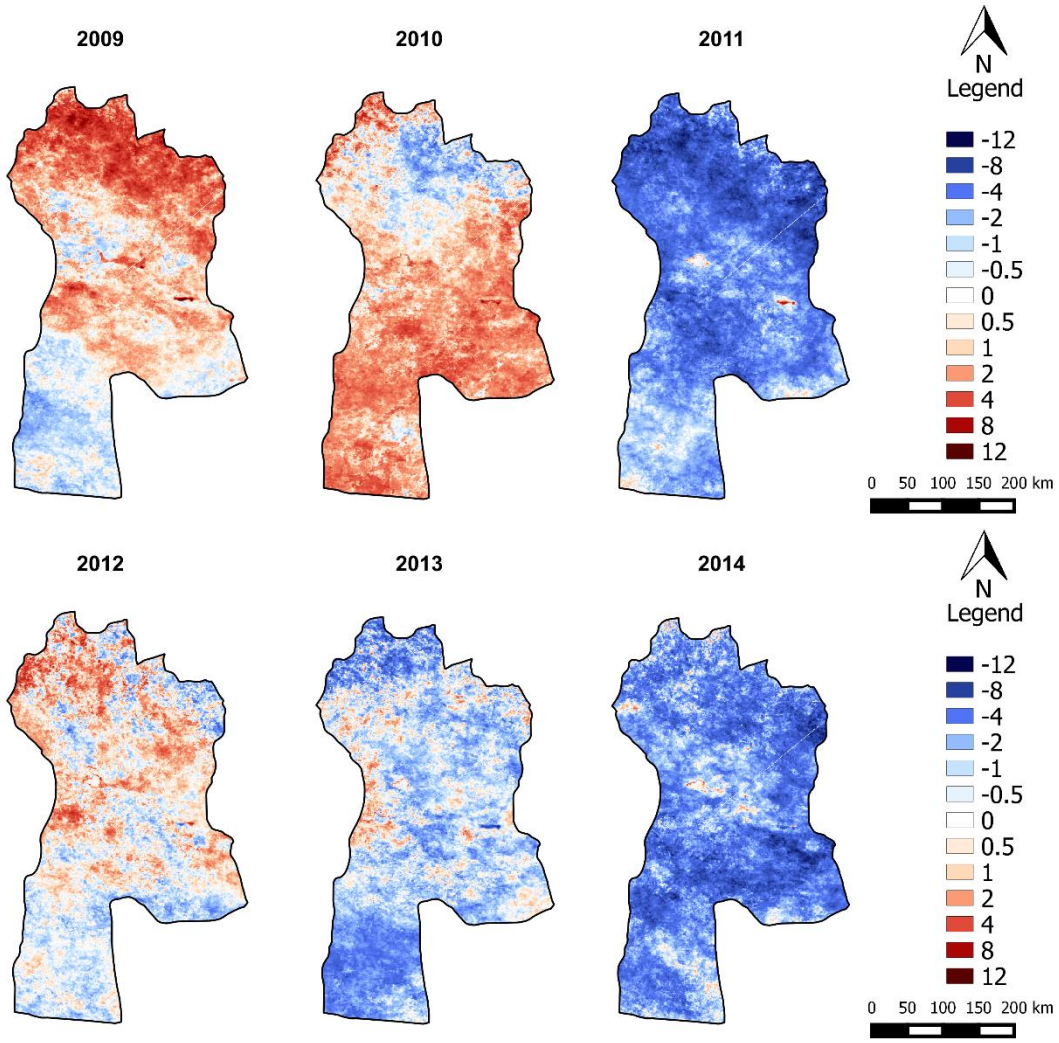


Figure 7.26 | Spatial distribution of LST anomalies of June for the period 2009 – 2014.

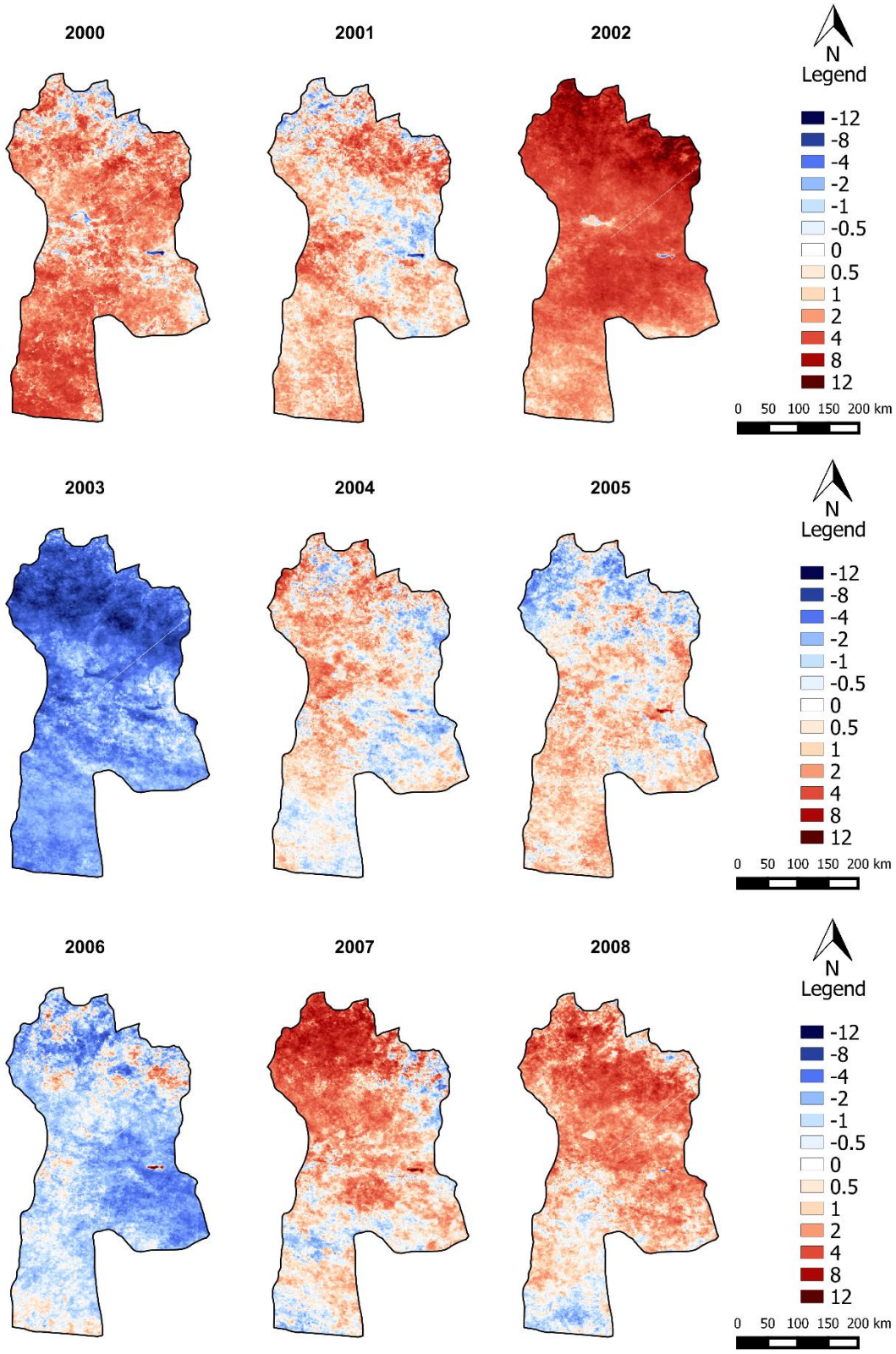


Figure 7.27 | Spatial distribution of LST anomalies of July for the period 2000 – 2008.

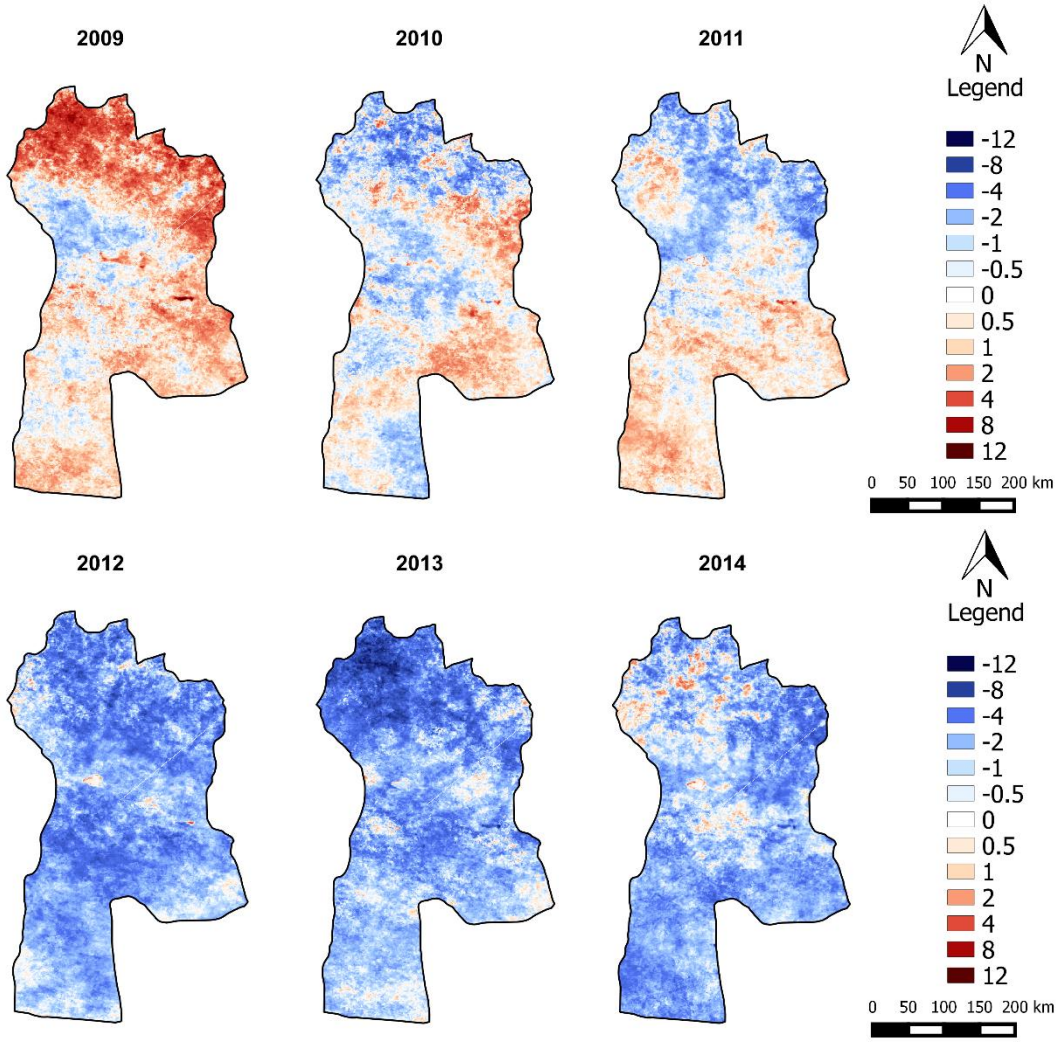


Figure 7.28 | Spatial distribution of LST anomalies of July for the period 2009 – 2014.

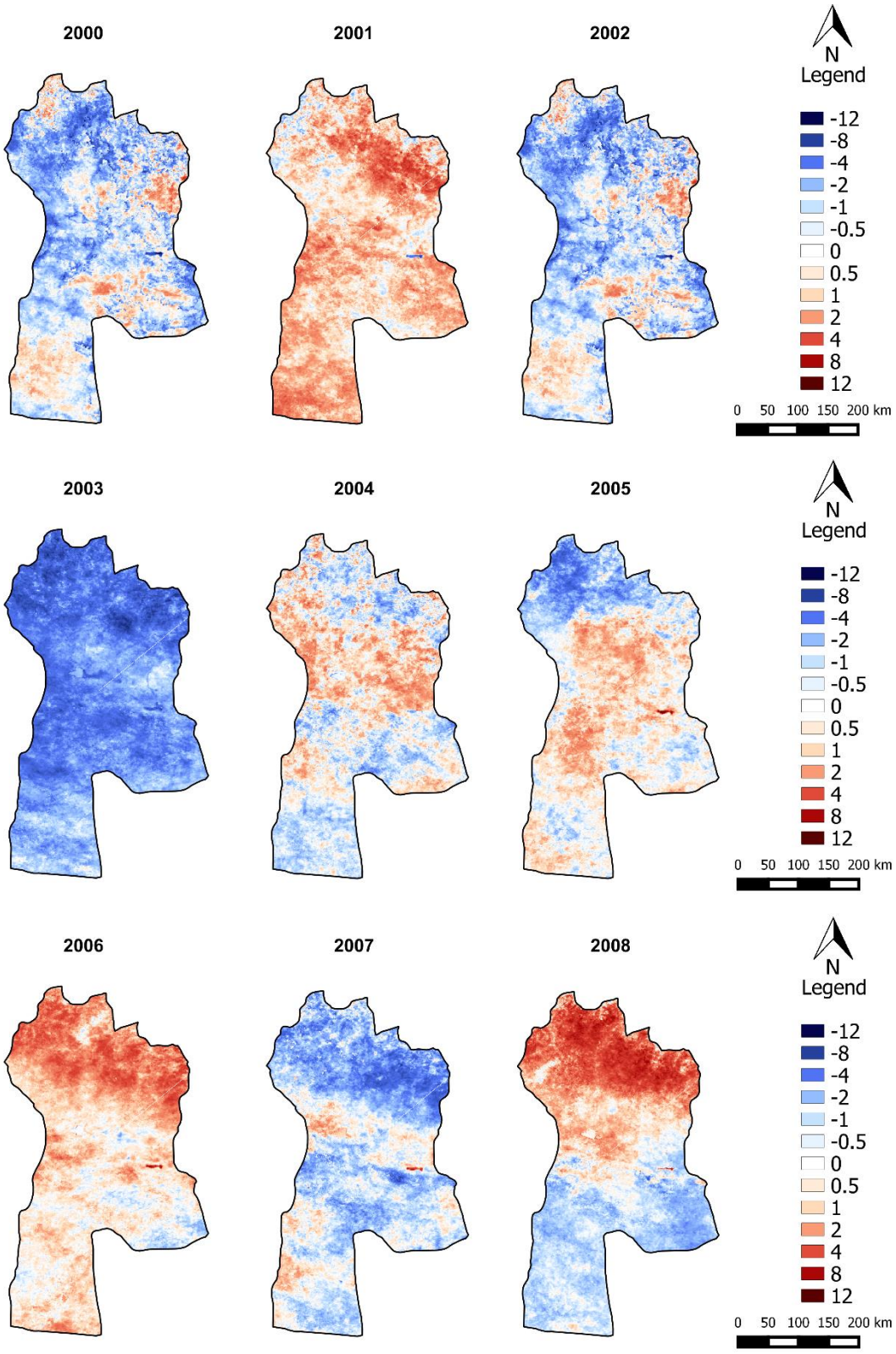


Figure 7.29 | Spatial distribution of LST anomalies of September for the period 2000 – 2008.

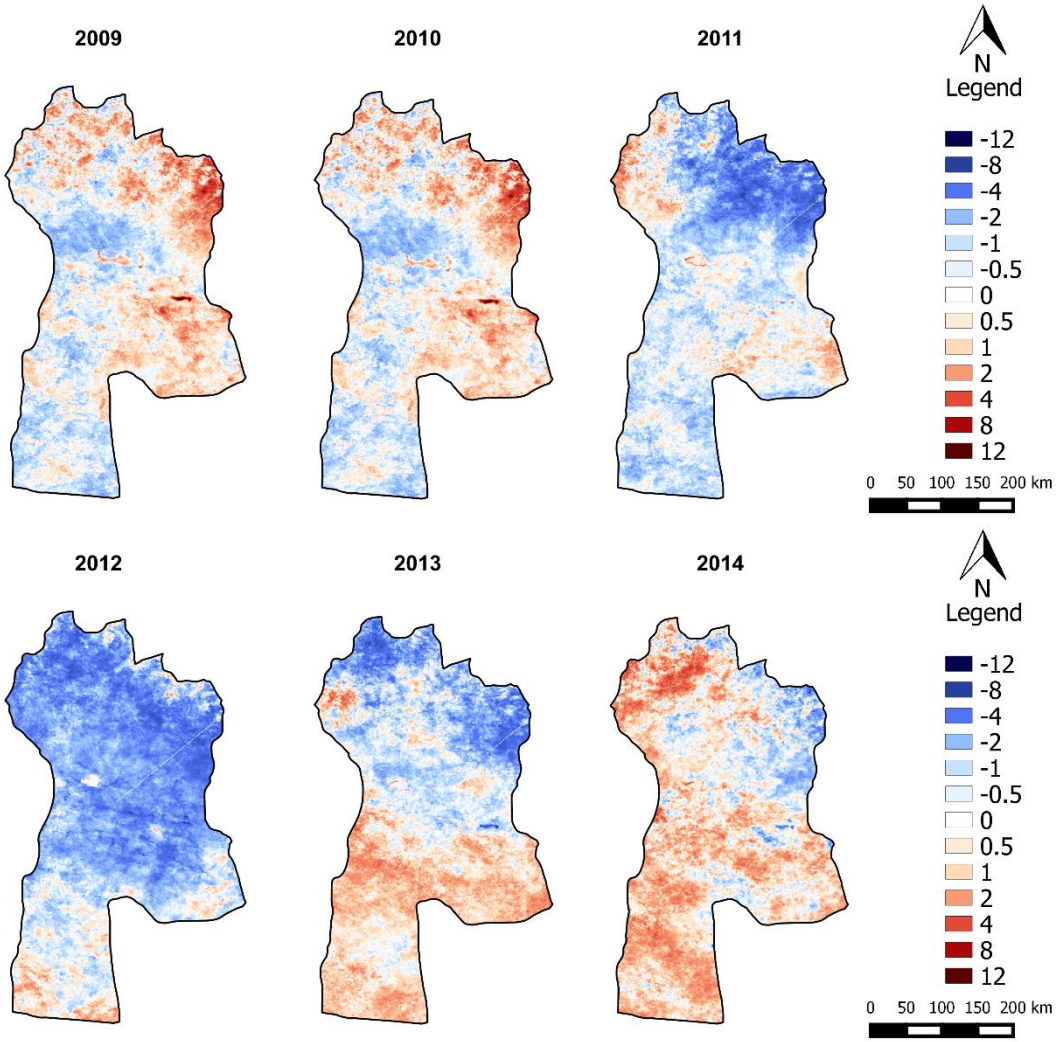


Figure 7.30 | Spatial distribution of LST anomalies of September for the period 2009 – 2014.

Appendix F – Monthly Precipitation anomalies

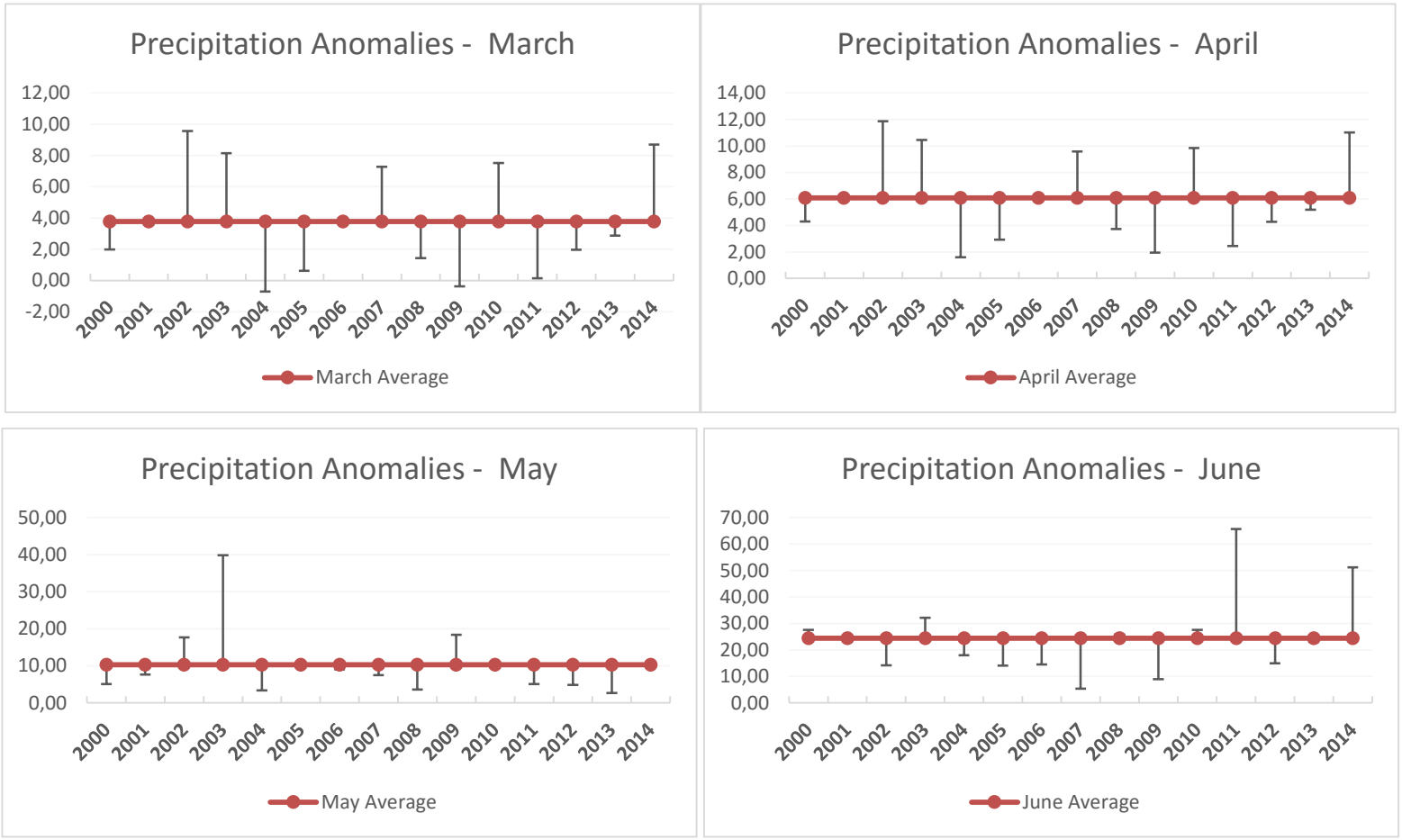


Figure 7.31 | Precipitation for the time series 2000-2014 (March – June) (mm)

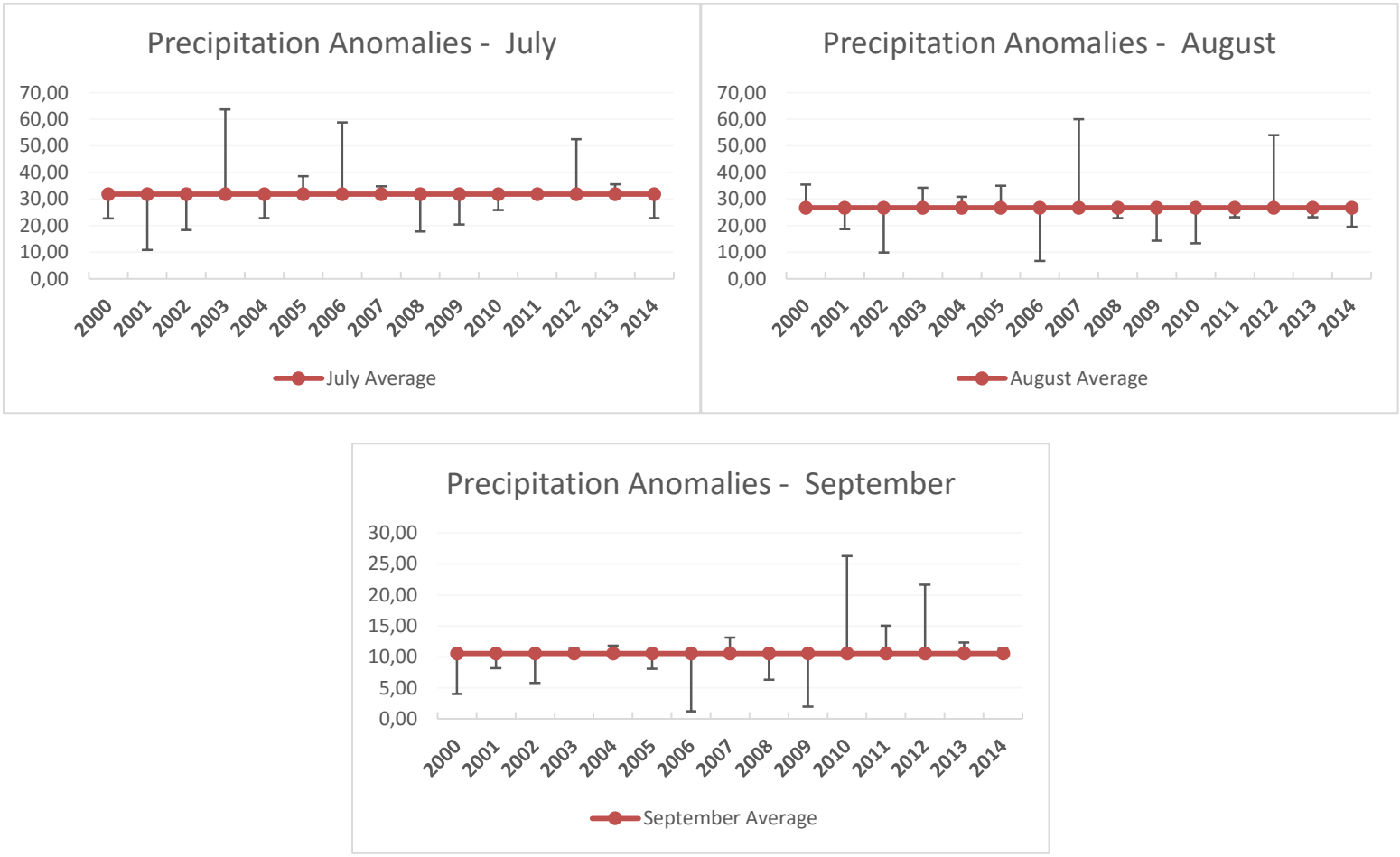


Figure 7.32 | Precipitation for the time series 2000-2014 (July – September) (mm)

Appendix G – Spatial Distribution of Precipitation anomalies

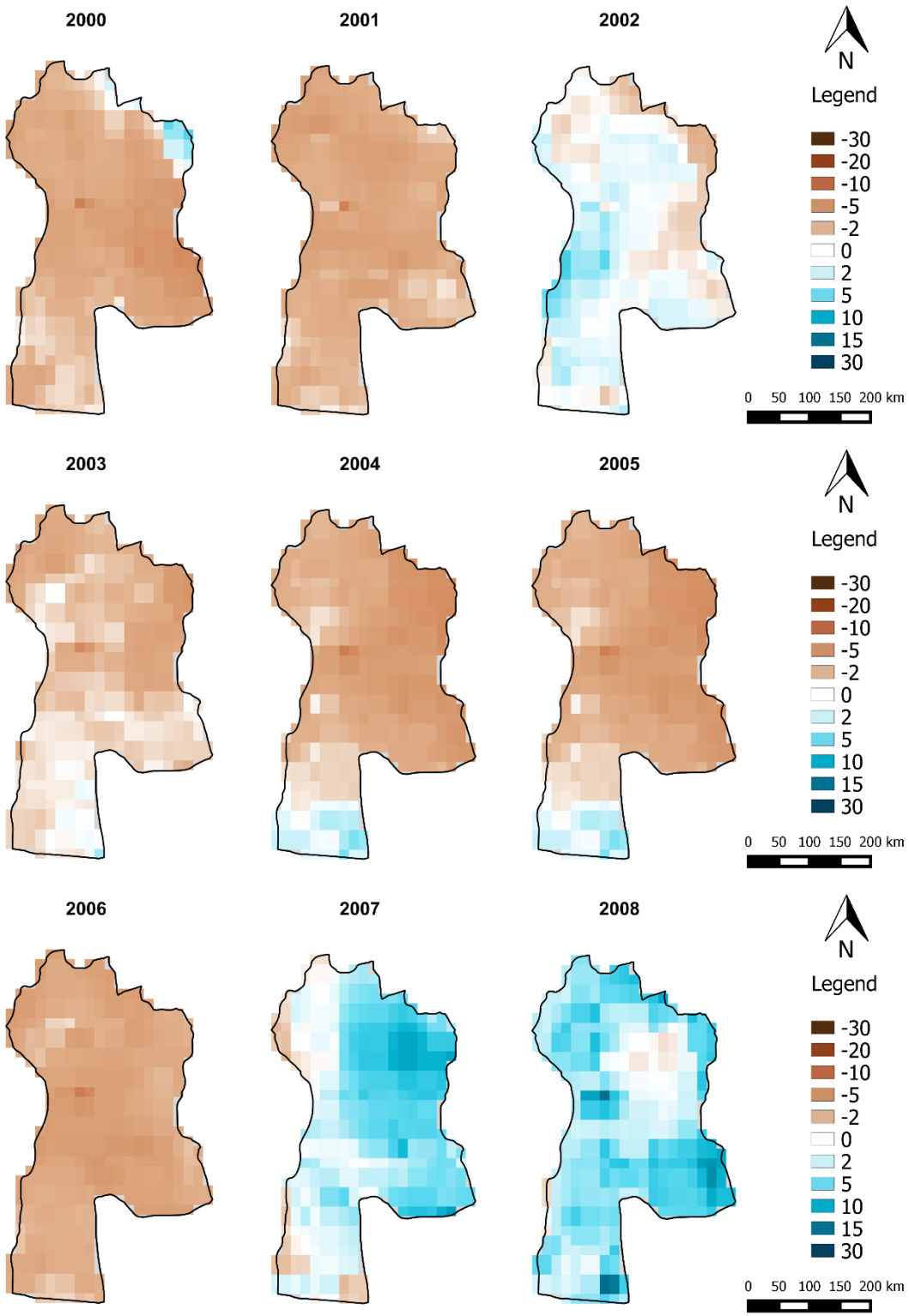


Figure 7.33 | Spatial distribution of precipitation anomalies of March for the period 2000 – 2008.

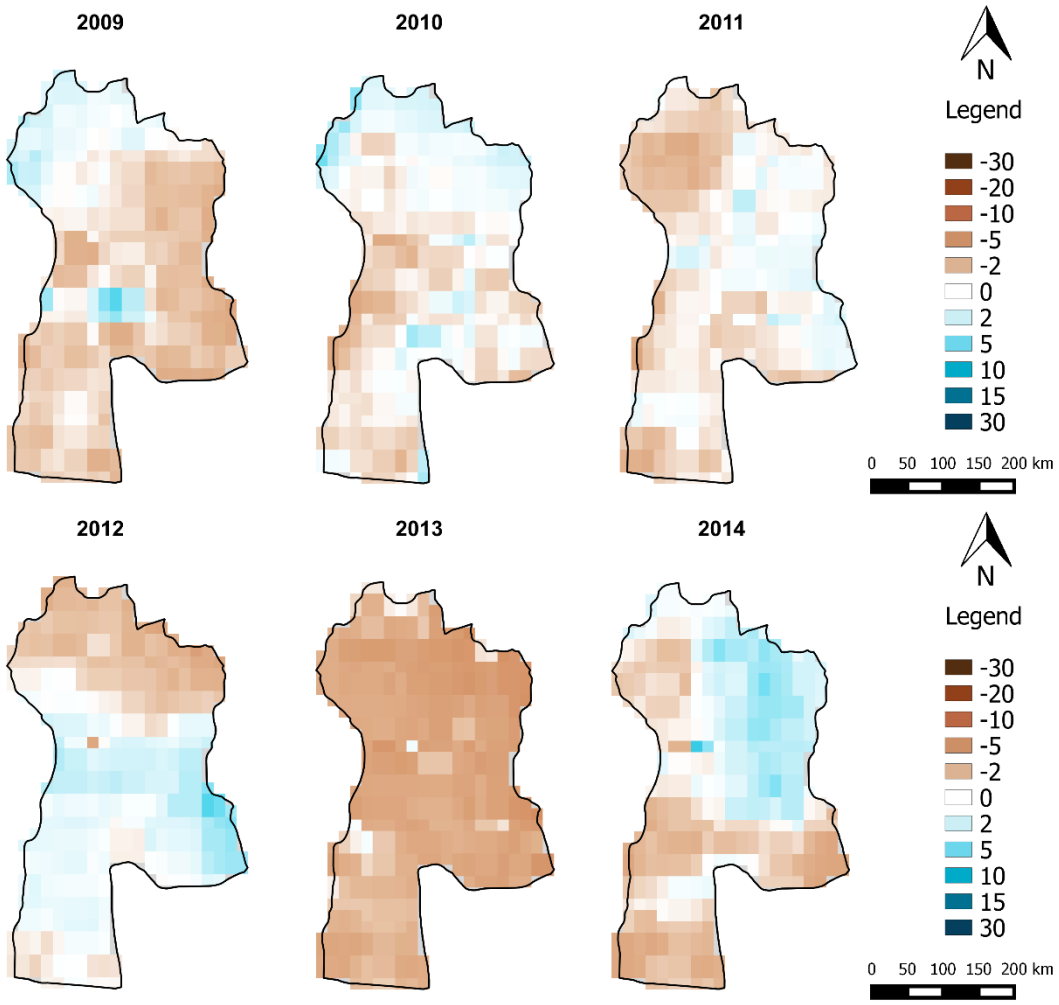


Figure 7.34 | Spatial distribution of precipitation anomalies of March for the period 2009 – 2014.

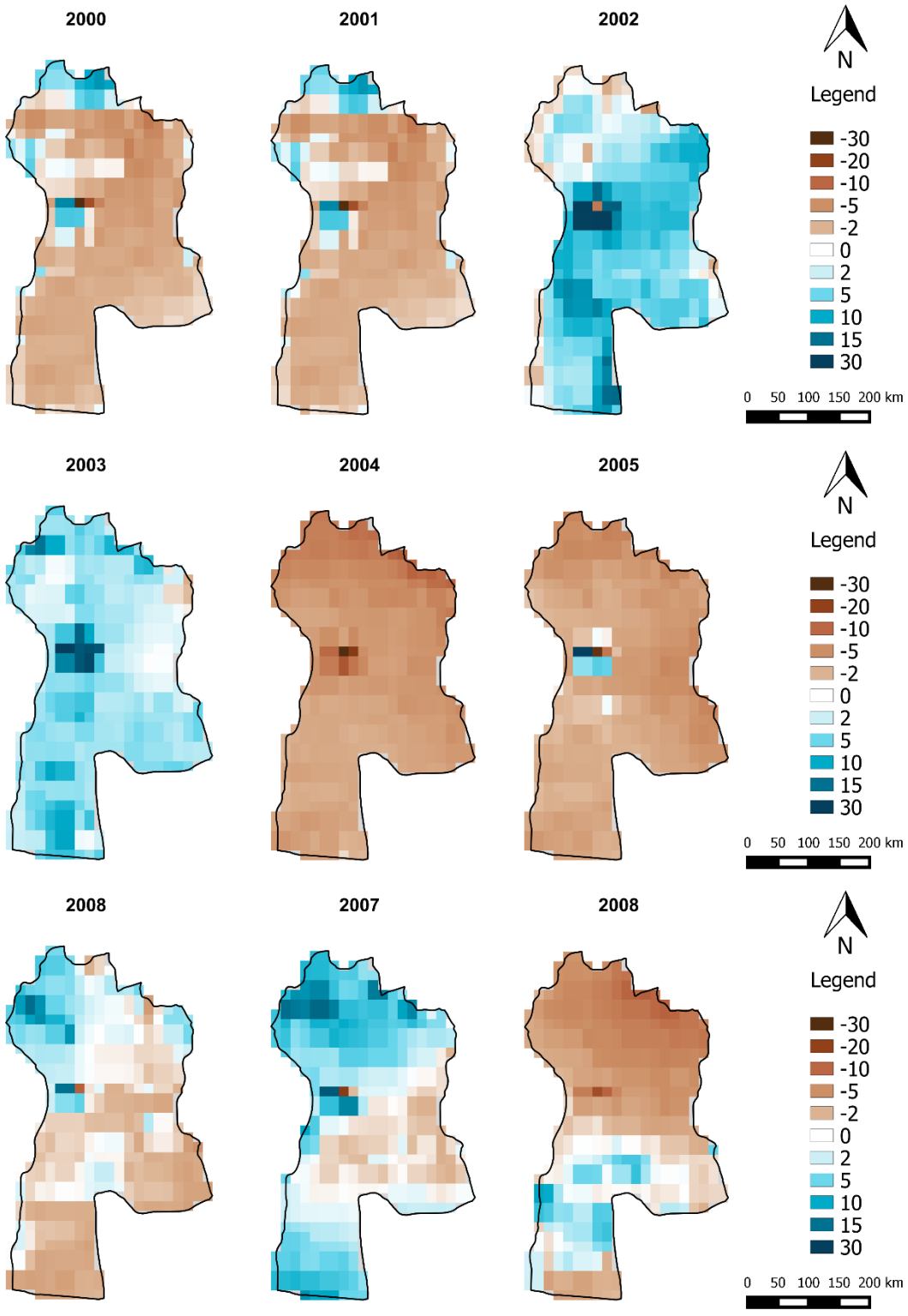


Figure 7.35 | Spatial distribution of precipitation anomalies of April for the period 2000 – 2008.

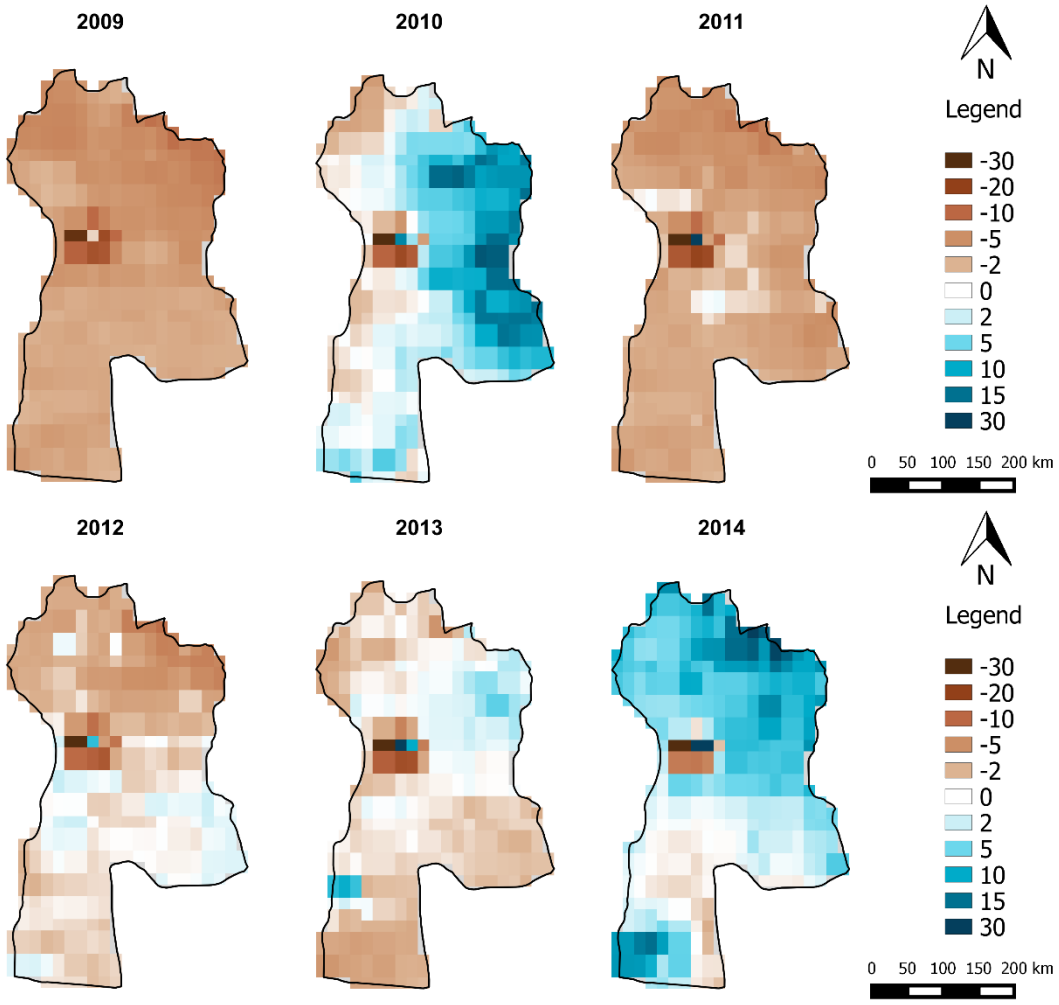


Figure 7.36 | Spatial distribution of precipitation anomalies of April for the period 2009 – 2014.

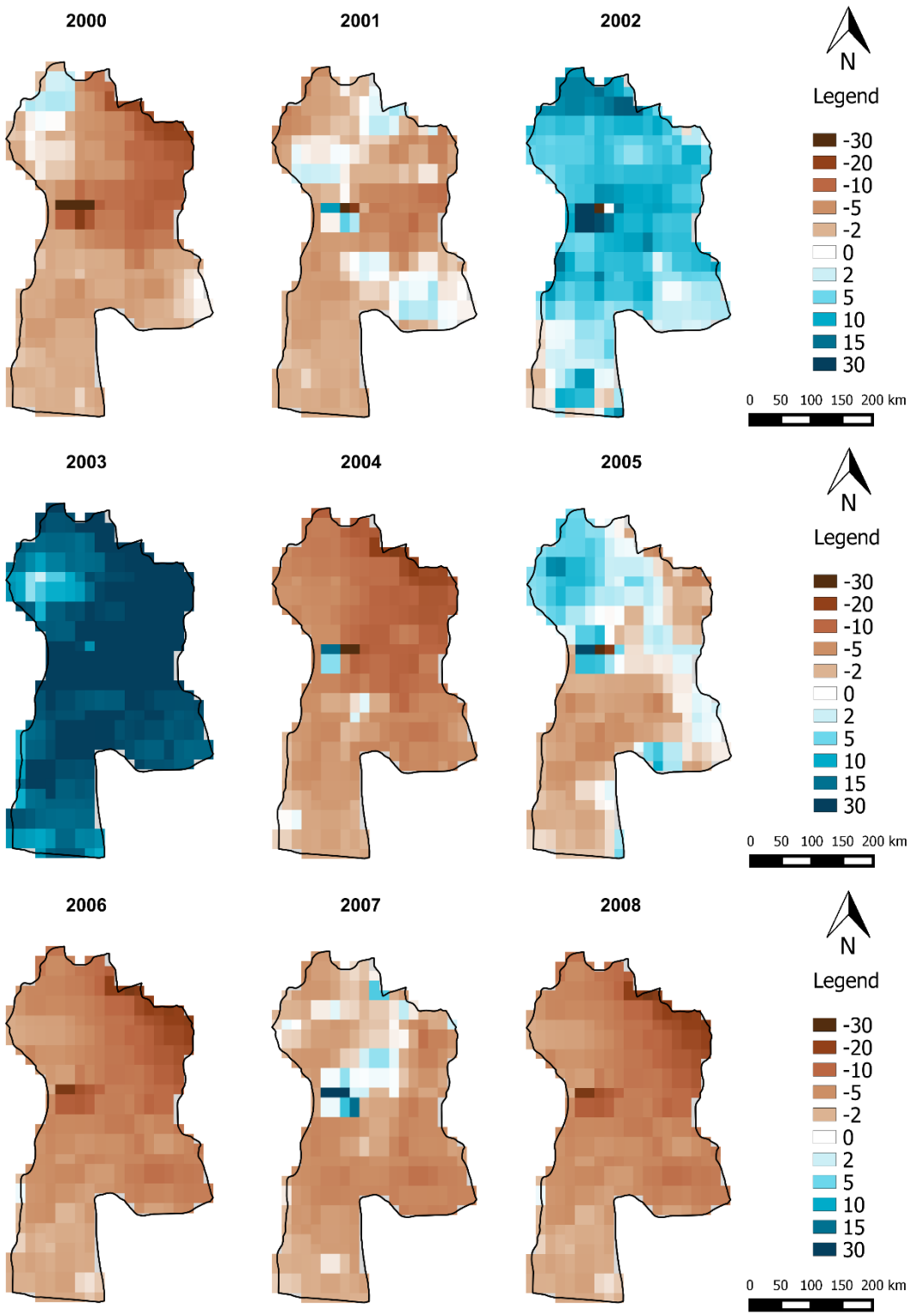


Figure 7.37 | Spatial distribution of precipitation anomalies of May for the period 2000 – 2008.

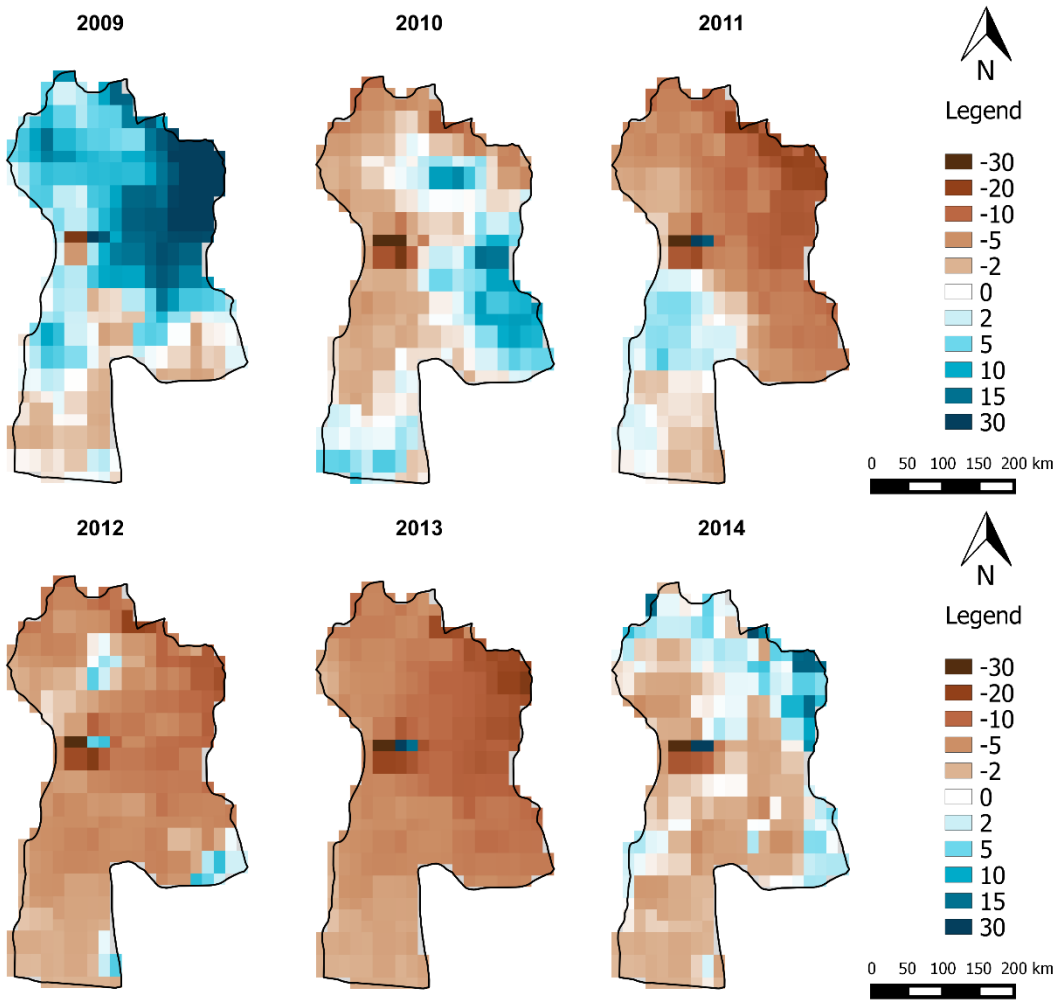


Figure 7.38 | Spatial distribution of precipitation anomalies of May for the period 2009 – 2014.

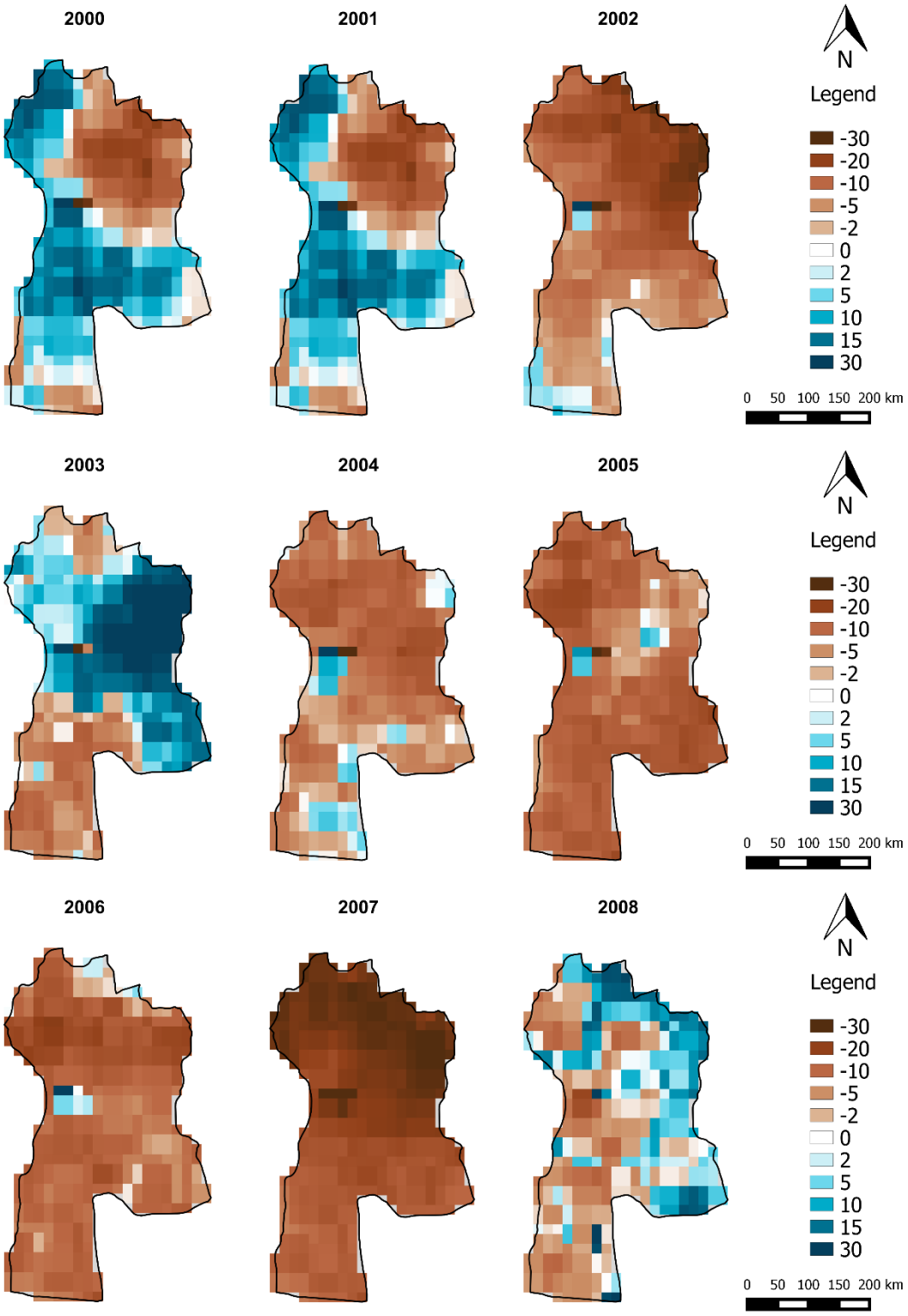


Figure 7.39 | Spatial distribution of precipitation anomalies of June for the period 2000 – 2008.

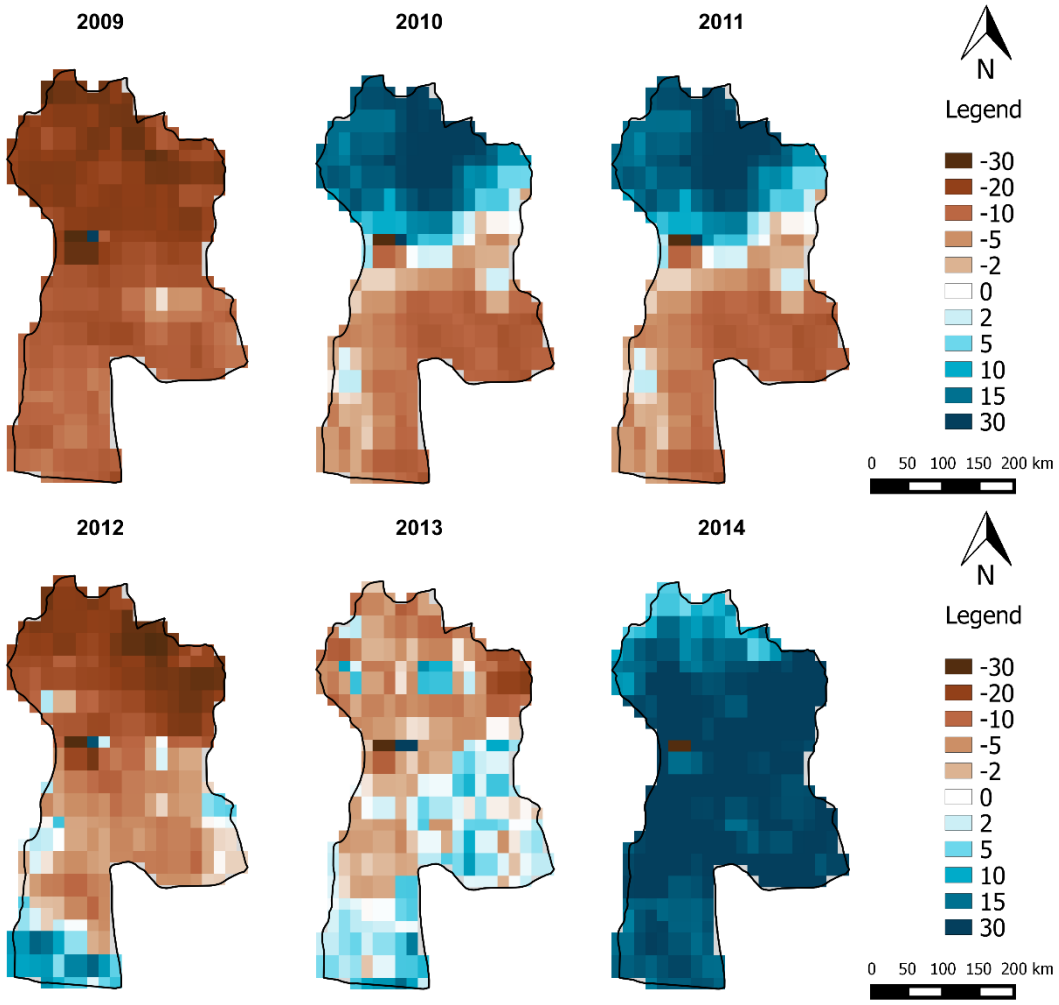


Figure 7.40 | Spatial distribution of precipitation anomalies of June for the period 2009 – 2014.

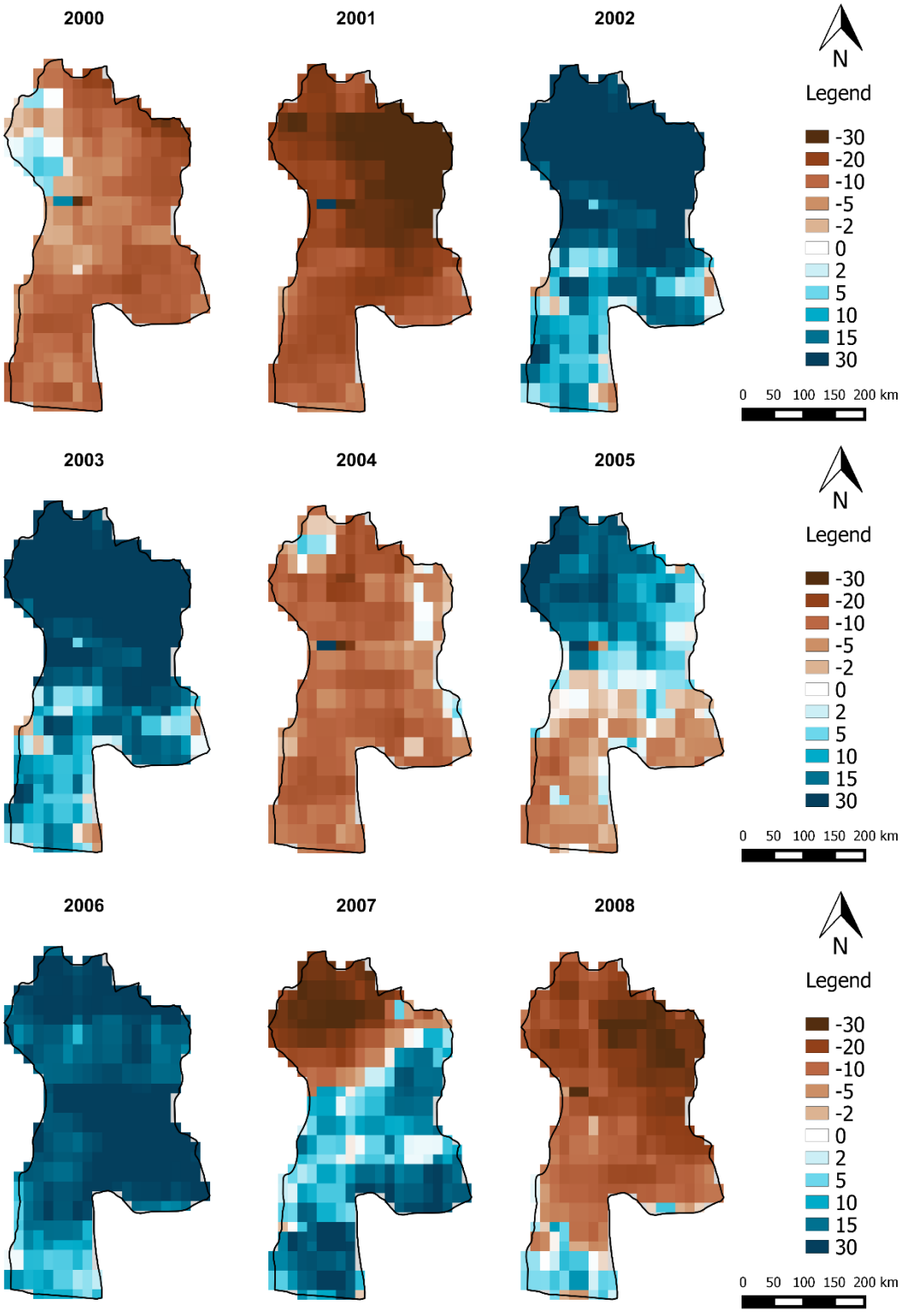


Figure 7.41 | Spatial distribution of precipitation anomalies of July for the period 2000 – 2008.

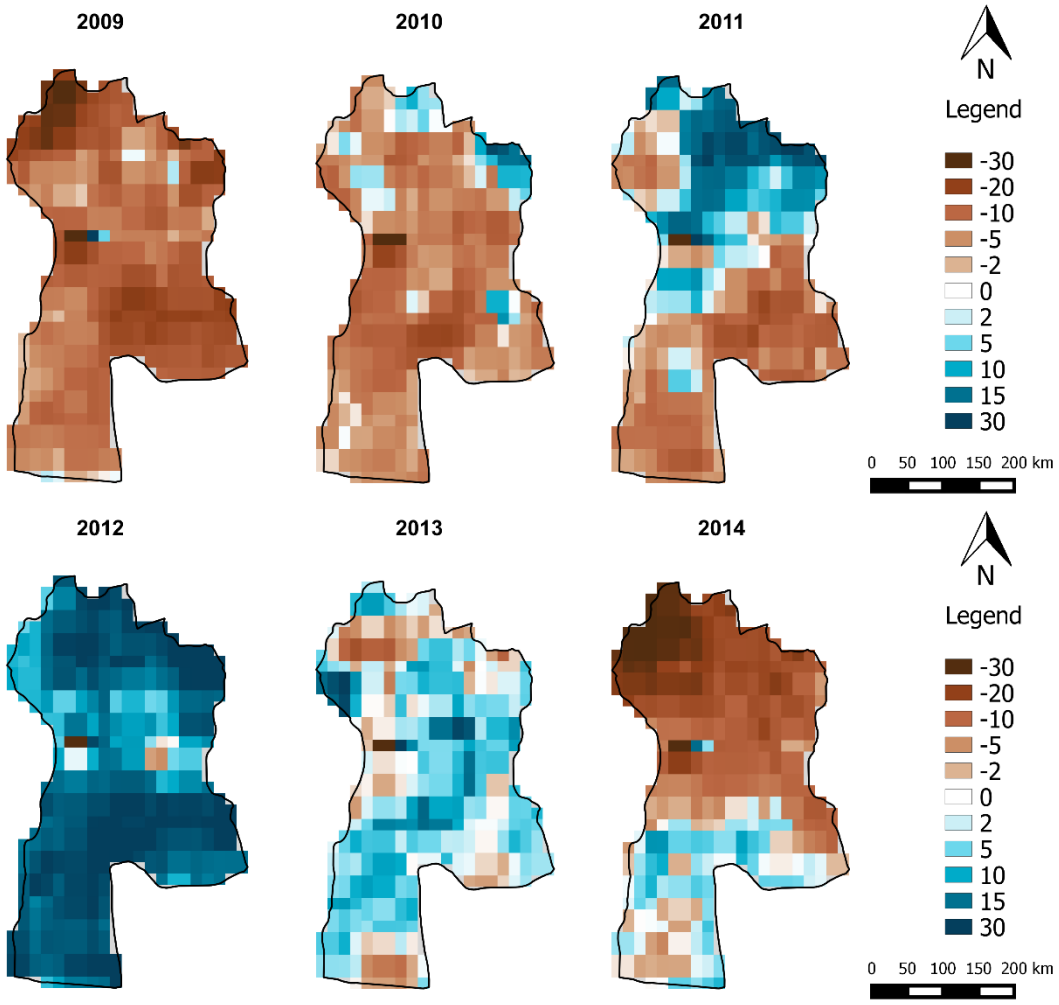


Figure 7.42 | Spatial distribution of precipitation anomalies of July for the period 2009 – 2014.

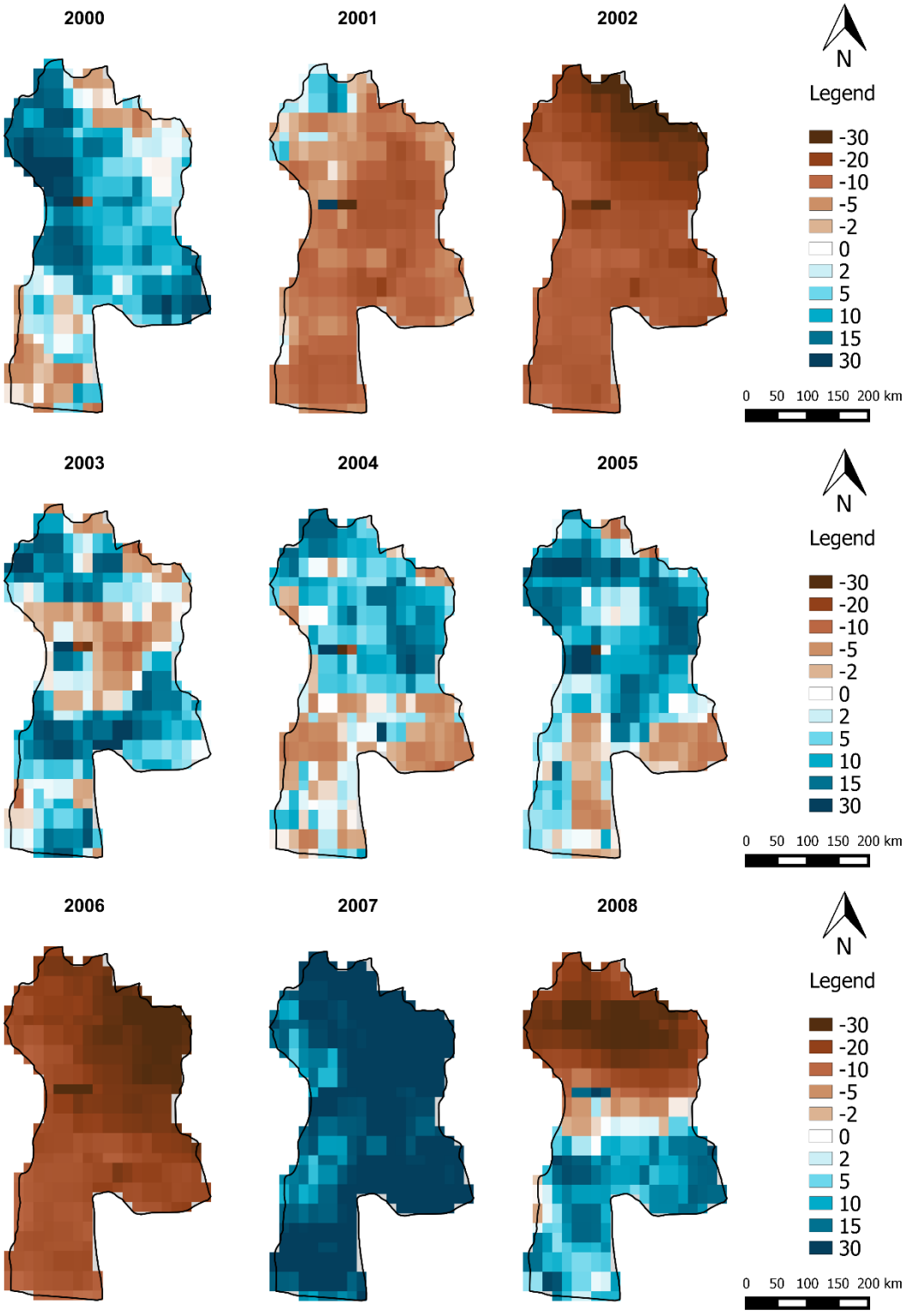


Figure 7.43 | Spatial distribution of precipitation anomalies of August for the period 2000 – 2008.

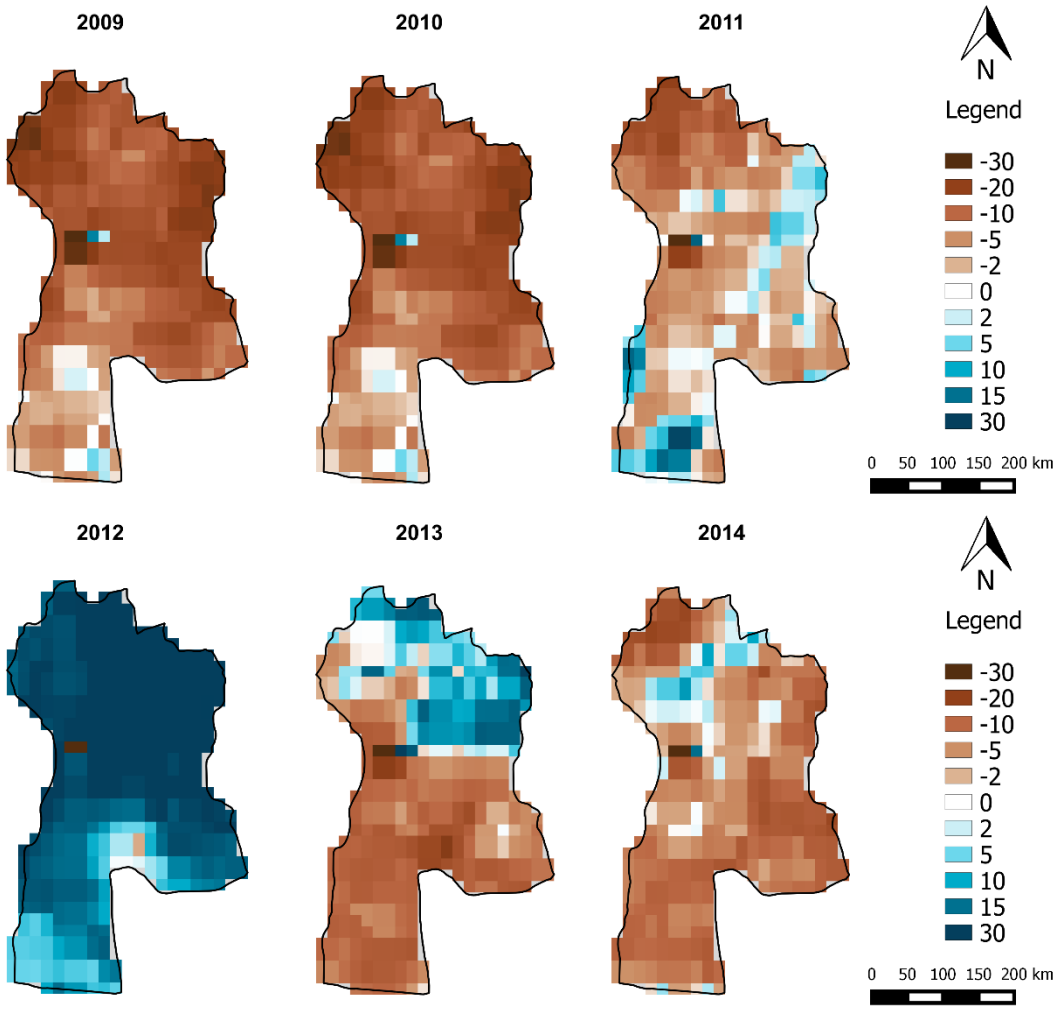


Figure 7.44 | Spatial distribution of precipitation anomalies of August for the period 2009 – 2014.

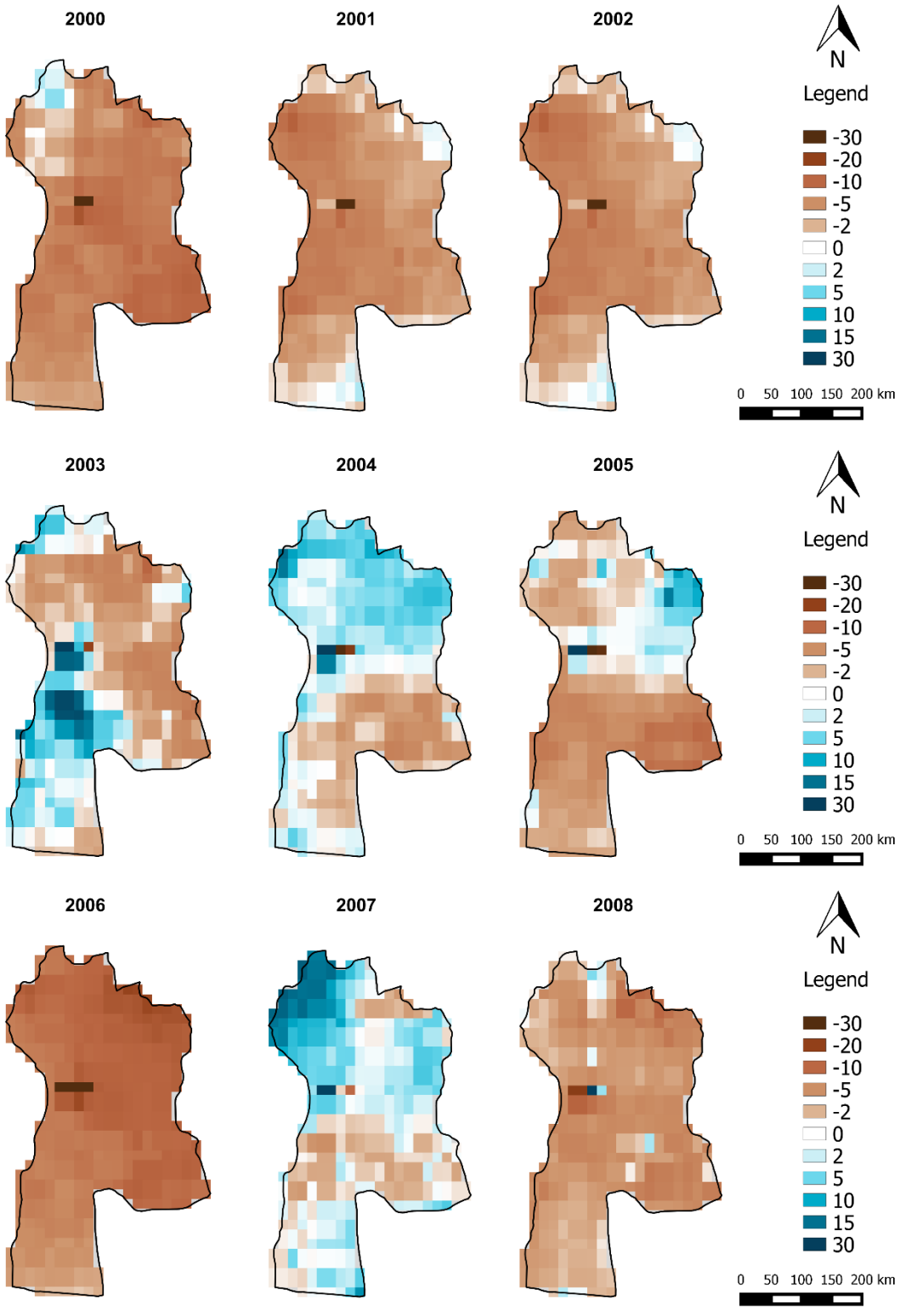


Figure 7.45 | Spatial distribution of precipitation anomalies of September for the period 2000 – 2008.

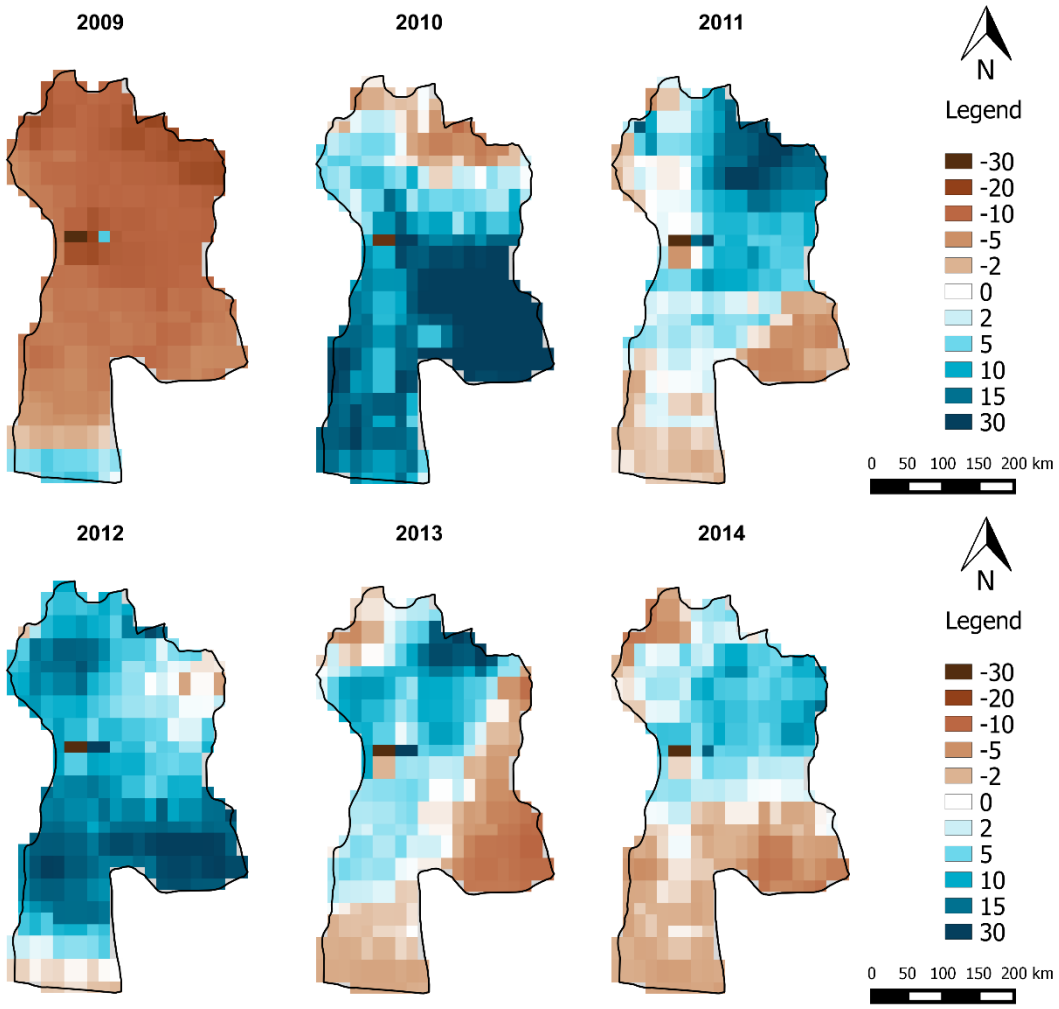


Figure 7.46 | Spatial distribution of precipitation anomalies of September for the period 2009 – 2014.

Appendix H – Programming for Everybody (Python) Certificate

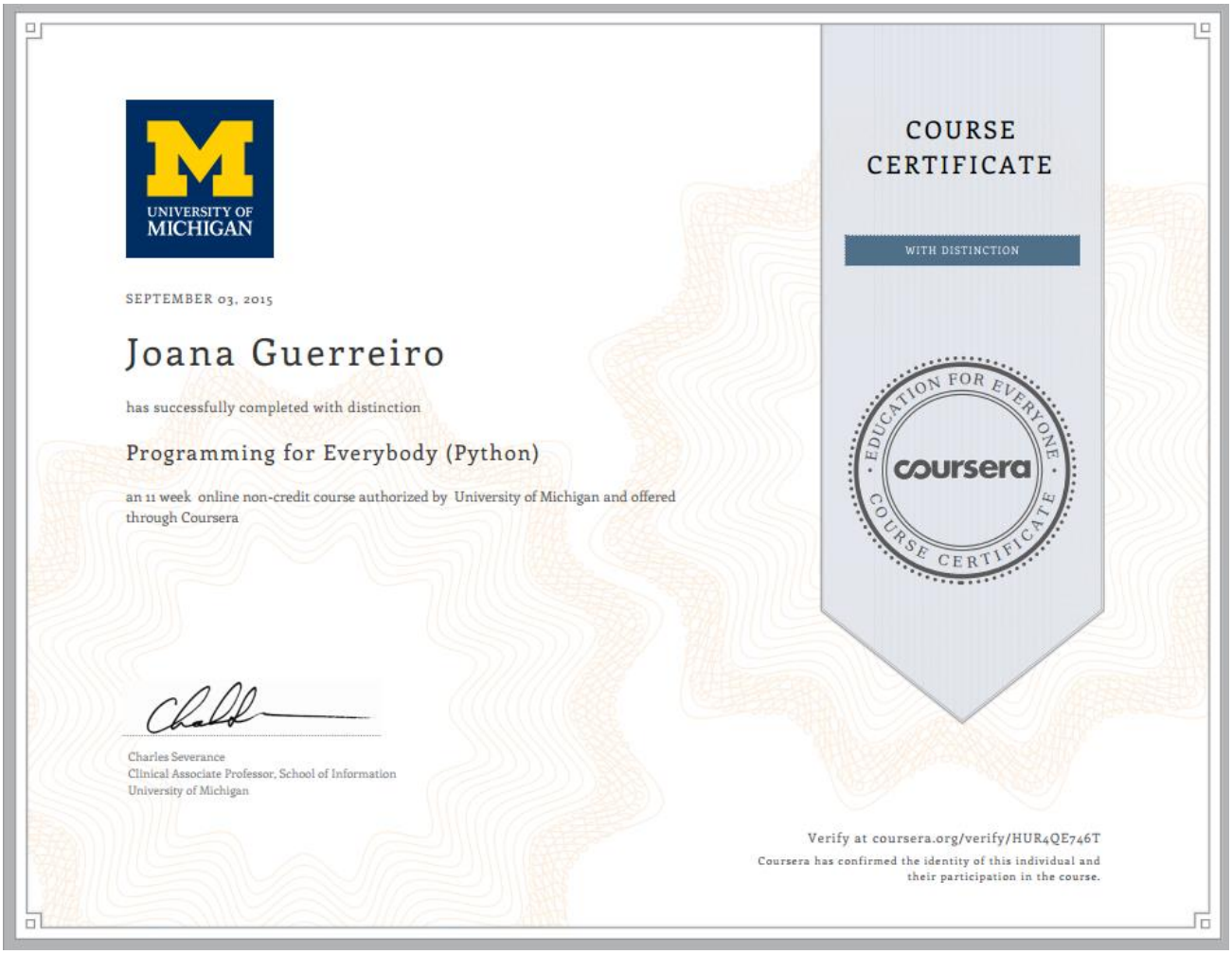


Figure 7.47 | Programming for Everybody (Python) by University of Michigan on Coursera.

Certificate earned on August 26, 2015. Further details on the certificate and course can be accessed through the URL: https://www.coursera.org/account/accomplishments/records/qGNj8PznekGDrCp